Published in final edited form as:

Chem Rev. 2013 November 13; 113(11): 8583–8603. doi:10.1021/cr400105n.

RNA Polymerase II Transcription Elongation Control

Jiannan Guo and David H. Price

Biochemistry Department, University of Iowa, Iowa City IA, 52242

1. Introduction

Regulation of gene expression is critical in determining cell identity, development and responses to the cellular environment. DNA is the inherited source of genetic information and regulation of gene expression starts with the selection of which genes will undergo transcription. RNA, the product of transcription, is then utilized to generate functional products, including being translated into protein or processed into functional RNA. In eukaryotes, protein coding genes are transcribed by RNA polymerase II (Pol II) into messenger RNAs (mRNA). These short-lived RNA species have a variety of characteristics and are extensively regulated from production to degradation.¹ With the assistance of methods such as those using microarrays and high-throughput sequencing, the scale and depth of Pol II transcription studies have exploded. The sheer volume and complexity of data from many sources have even triggered a call for careful rethinking of the methods used for analysis and interpretation.² It is doubtless, though, that regulation of transcription critically affects gene expression and thus cell state and cellular identity.³

Pol II transcription starts with the assembly of a pre-initiation complex (PIC) with general transcription factors (GTFs) that recognize DNA sequence elements around the promoter and recruit Pol II.⁴ This process also requires the multi-subunit Mediator complex that could be viewed as a platform for transcription.⁵ In the PIC, the two strands of DNA are separated and the template strand migrates into the active center of Pol II, thereby allowing the synthesis of RNA from the transcription start site (TSS).⁶

Although initiation could be viewed as the “on” switch for Pol II, much of mRNA production is regulated at the elongation phase.⁷ Pioneering studies on *MYC*⁸, *HIV*⁹, and *HSP70*¹⁰ transcription have indicated that Pol II can be transcriptionally engaged in the 5' end of genes without generating full-length mRNA prior to induction. Genome wide analyses showed that a large fraction of human and *Drosophila* genes have poised Pol II about 50 nt downstream of the transcription start site (TSS).¹¹ Under various activation conditions, Pol II is released from promoter proximal positions to produce full length transcripts and subsequently increase mRNA level.¹² The factor required to trigger Pol II to enter productive elongation is P-TEFb.¹³ Productive elongation has a high elongation rate that ranges from 1.1 to 4.3 kb/min as measured by many different methods.¹⁴ During productive elongation the RNA is co-transcriptionally spliced and polyadenylated to generate mature mRNAs.¹⁵ Mirroring the dramatic differences in properties, productive

elongation complexes have significantly different protein compositions than early elongation complexes.¹⁶

Transcription termination is crucial for recycling Pol II after a round of transcription and globally releasing Pol II from chromatin prior to cell division.¹⁷ It also helps to prevent interference of promoter function by transcription from neighboring genes.¹⁸ In metazoans, Pol II termination downstream of the 3' end of almost all protein coding genes requires a functional Poly(A) signal and is always coupled with 3' end processing.¹⁹ Because termination is the end of transcription elongation and by definition is a very transient state, it has been notoriously difficult to study, especially in vivo.²⁰

The steps in transcription have been traditionally studied individually in great depth using specific genes. The development of new technologies has allowed transcription to be viewed and studied on a global scale. This review discusses the bird's eye view of Pol II transcription in the genome as well as insights provided by detailed mechanistic studies. Recent studies are emphasized, but initial discoveries are also described to provide a historical perspective. We mostly focus on metazoan systems, although some studies from yeast are also described for comparative purposes. Our goal is to cover topics in multiple levels so that beginning scientists as well as experienced researchers will find the review useful.

2. Pre-elongation transcriptional events

2.1 Promoter access within chromatin

By default eukaryotic DNA is covered by nucleosomes limiting access of the transcription machinery to promoters (Figure 1). Although nucleosomes have limited intrinsic preferences for specific DNA sequences²¹ this cannot explain how most promoters are made available for transcription. It is generally accepted that specific DNA binding factors recruit chromatin remodelers that regulate access to promoters. As described later, once the transcription machinery is engaged it may help to maintain the promoter in an open configuration. A complete discussion of how the block by repressive nucleosomes is overcome is beyond the scope of this review and readers are referred to other recent reviews of the subject.²²

2.2 Initiation and general transcription factors

Pol II cannot initiate efficiently or in a site-specific manner without the aid of GTFs which have been identified and extensively characterized biochemically as reviewed (Figure 1).^{4,23} GTFs and Pol II together form a pre-initiation complex (PIC) on core promoter sequences.²⁴ The multi-subunit TFIID complex recognizes core promoter elements if they are present including the TATA box, initiator, and at least in *Drosophila*, two downstream promoter elements MTE and DPE.²⁵ It should be noted that the majority of promoters in vertebrates are comprised of loosely defined CpG islands,²⁶ which appear to be associated with bidirectionality of initiation (see section 4.2). Identification and characterizations of the GTFs were carried out in vitro using promoters containing conserved core elements. Although not required in vitro, TFIIA stabilizes the TFIID-DNA interaction.²⁷ TFIIB directly makes contact with Pol II²⁸ and Pol II interacts tightly with TFIIF, which can stabilize the TFIIB-Pol II interaction.²⁹ TFIIF also stimulates the recruitment of TFIIF,³⁰

which is a multi-function complex with three enzymatic activities that affect transcription. Two are ATP-dependent DNA helicases, XPB and XPD, and one is a kinase, CDK7, that phosphorylates the carboxyl terminal domain (CTD) of the large subunit of Pol II. The XPB 3'-5' helicase activity unwinds the DNA to form a transcription bubble with the template strand positioned near the active site of the polymerase.³¹

The order of entry of TFIIB, TFIIE, TFIIIF and TFIIH and Pol II into the PIC is not universally agreed upon. Early models proposed either a sequential assembly of the factors and Pol II or the recruitment of a Pol II holoenzyme containing a number of factor to a minimal DNA recognition complex.^{24a} In the human system a complex between the DNA, TFIID (with TFIIA), TFIIB has been observed and this recruits Pol II, TFIIIF and finally TFIIE and TFIIH.³² However, reconstitution of the yeast PIC containing 31 proteins identified an early intermediate containing DNA, the TBP subunit of TFIID, TFIIE and TFIIH that could recruit TFIIB, Pol II and TFIIIF.^{27b} Because there are a large number of interactions between the polymerase, factors and DNA it is not surprising that there may be different orders of factor entry into the complete PIC and this may be promoter specific.³³ Regardless of the order of assembly a fully formed PIC will initiate and may carry out a number of rounds of abortive initiation in which short nascent transcripts are repeatedly synthesized.³⁴ Once the transcript attains a length of about 14 nt the polymerase exits initiation and enters the elongation phase.³⁴

2.3 Mediator and Gdown1

GTFs are required for Pol II initiation on every promoter, regardless of the strength of the promoter. The regulation by activators and repressors are conveyed through a conserved multi-subunit complex termed the Mediator.^{5,35} The human Mediator contains about 30 subunits and is 1.2 MDa in size, making the PIC approximately 3.5 MDa.³⁶ Known to interact with Pol II and several GTFs, Mediator is viewed as a platform that responds to different activators for regulating transcription in different ways.³⁷ During large scale purification of Pol II from calf thymus the Gnatt lab found two forms of Pol II that differed in subunit composition.³⁸ One form, Pol II(G), contained a 13th subunit called Gdown1 that is encoded by the *POLR2M* gene. The presence of Gdown1 repressed the effect of activators on Pol II transcription in vitro and this repression was alleviated by Mediator.³⁸ In this way Pol II(G) is seen to be the Mediator responsive form of Pol II. It was later shown that Gdown1 competes with TFIIIF for interaction with the RPB1 and RPB5 subunits of Pol II.³⁹ In vitro transcription using a defined system has shown that TFIIIF's effect could be achieved through transient interaction with the PIC instead of a stable contact.²⁹ The dynamic sequence of contacts among Pol II, Gdown1 and TFIIIF during initiation remains to be elucidated.

3. Methods for studying Pol II transcription

Pol II transcription machinery involves a large number of factors to regulate initiation and elongation. A wide variety of assays has been developed to study the functions of these factors. The first assays developed simply followed NTP incorporation by Pol II using templates comprising genomic ssDNA or dsDNA with single stranded ends that allow initiation of Pol II without factors. The GTFs were discovered and identified using

fractionated nuclear extracts to reconstitute accurate initiation of transcription on dsDNA templates containing promoters.⁴⁰ Elongation factors were then discovered by their effects on isolated elongation complexes using immobilized templates.⁴¹ In vitro assays have gotten more elaborate with the incorporation of chromatinized templates and the inclusion of DNA-binding transcription factor binding sites.⁴² More recently, global methods have been developed to analyze the precise genome-wide occupancy of the transcription machinery, as well as the transcriptome in cells.²

3.1 In vitro methods

In vitro transcription systems can be used to elucidate detailed mechanistic functions of individual factors, combinations of factors, or activities in crude extracts. In the defined system, purified Pol II with GTFs (optionally with Mediator and activators) are assembled onto a promoter containing DNA template and allowed to initiate by the addition of NTPs.^{29,38,39b,43} Radiolabeling one of the nucleotides (usually α -³²P-CTP) provides a means to measure newly synthesized RNA. For the subsequent quantifications and comparisons, transcription is often allowed to elongate to a fixed point on the template. This could be the end of the template if the polymerases are allowed to run off or at specific positions dictated by the combinations of NTPs present in the reaction.^{29,43–44} These purified systems can produce definitive results on the functions of purified factors.

An easier and quicker way to initiate transcription in vitro is through the use of nuclear extracts.⁴⁵ While sacrificing the knowledge of the exact composition of PIC, this method efficiently generates elongation complexes (ECs). The ECs can be isolated using immobilized templates, which enables further analyses of elongation.^{41b,46} In addition to analyzing the properties of the elongation process and the stability of the nascent transcript,¹⁷ the composition and conformation of the complex can also be analyzed using an electrophoretic mobility shift assay (EMSA).^{39a,47} Because the elongation complex is stable to 1.6 M KCl⁴⁶ it can be stripped of factors and the migration of the elongation complex detected by the radioactive nascent transcript. Elongation complexes that associate with added factors have altered mobility.^{47a}

3.2 Global methods

Chromatin immunoprecipitation (ChIP) captures protein-DNA interactions.⁴⁸ ChIP involves crosslinking proteins to the DNA, shearing genomic DNA to small fragments and isolating the protein of interest with antibodies. The DNA pulled-down with the protein can be identified and aligned to the genome and the resulting distribution reflects the genomic occupancy of the targeted protein. DNA binding transcription factors, Pol II, and factors that associate with Pol II are frequently studied by ChIP. The DNAs pulled down can be analyzed by PCR to examine specific regions or by microarray (ChIP on chip) or high-throughput sequencing (ChIP-Seq) for better resolution on a genome-wide scale (Figure 2A). Antibodies are available for most known transcription factors and as long as they work in immunoprecipitations, the technique can analyze the genome occupancy of many of the proteins that are bound to DNA. Because formaldehyde crosslinks proteins not only to DNA, but also to other proteins, even factors that are only peripherally associated with DNA can be mapped.

There are limitations to ChIP that should be considered. Because many antibodies recognize only one or predominately one epitope, masking of the epitope by bound factors or by post-translational modifications can potentially lead to missing some bound regions. Off target interactions of antibodies can also lead to the reverse situation with bound regions being incorrectly attributed to the factor of interest. Other issues to consider concern quantitative analysis of the resulting data. ChIP-seq datasets can be considered reasonably good for determining relative signal strength across any given dataset and this can be used to determine the differences and similarities between the distributions of two different factors. Absolute quantification of a single factor at any specific site from cells under different conditions is very difficult to achieve with accuracy. Part of this difficulty comes from the large number of non-specific fragments that are brought down in each immunoprecipitation. These fragments are sequenced and the resulting reads are usually greater than 95% of the total reads. In fact, the specific reads (those resulting from appropriate pull down of DNA bound by the protein of interest) are usually between 1 and 5% and in cases of bad datasets significantly less. Because of the variations in background reads between datasets, normalizing to the total number of reads does not allow for quantitative comparisons. The quality of datasets can be quickly determined by viewing the data using the UCSC Genome Browser.⁴⁹ Non-specific background is visible as a low, random distribution of signal between the peaks that reflect enriched occupancy.

Whilst ChIP-Seq provides valuable insights on the factor's distribution, RNA-Seq is a powerful complement that monitors global RNA levels (Figure 2B). To determine steady-state RNA levels, total RNA can be purified from cells and different fractions of the RNA (for example, Poly(A) plus mRNA) can be isolated before being reverse transcribed and sequenced.⁵⁰ The method can provide a very deep, strand specific view of the RNAs present in the cell. Of course, just as the methods from which it evolved, Northern Blotting and cDNA microarrays, RNA-seq does not directly provide a readout of transcription, since RNA processing events and mRNA stability impact the population of RNAs at any given time in a cell. Factors binding to the RNA can also be mapped by HITS-CLIP⁵¹ (high-throughput sequencing of RNA isolated by crosslinking immunoprecipitation) or PAR-CLIP⁵² for higher sensitivity (photoactivatable-ribonucleoside-enhanced crosslinking and immunoprecipitation).

Two newer methods in which nascent transcripts are sequenced give a strand-specific picture of active transcription. GRO-Seq (Global Run-On sequencing) protocol starts with a nuclear run-on assay in which nascent transcripts are extended and labeled under conditions that inhibit initiation (Figure 2C).⁵³ Sequencing the extended RNA yields information about the position of transcription complexes. The run-on reaction usually includes Sarkosyl which removes factors that impede elongation and, therefore, detects all engaged complexes. Comparing GRO-Seq data obtained with and without the use of Sarkosyl can be used to determine the positions of paused Pol II globally.⁵⁴ A modified protocol, PRO-Seq, allows single base resolution by carrying out reactions with each of the 4 biotinylated NTPs individually.⁵⁵ Another method, NET-Seq (Native Elongating Transcript sequencing), begins with a Pol II immunoprecipitation step before the RNA is isolated and sequenced (Figure 2D). It surveys nascent transcripts being generated and the data can be used to map

elongating Pol II to nucleotide resolution.⁵⁶ Data generated by both GRO-Seq and NET-Seq include only signals from elongation complexes not PICs.

4. Promoter proximal paused Pol II

4.1 The paused state allows regulation of elongation

Elongation in vitro by most RNA polymerases, especially Pol II, is characterized by the pattern of transcripts generated due to pausing and arrest. Pausing is practically defined as a temporary halt to elongation at a specific site along the template. The polymerase stays engaged, but requires an extended time to resume elongation and leave the site. If the polymerase never leaves the pause and does not terminate it is said to be arrested. At the maximum rate of 4 kb/min Pol II takes on average 15 milliseconds for each nucleotide addition. There is a wide range of dwell times at specific pause sites in the absence of factors with the strongest being more than 4 orders of magnitude longer (>10 minutes). The mechanisms of pausing and arrest involve conformational changes in the polymerase, template and RNA. This topic was thoroughly discussed in a recent review⁵⁷ and will not be covered in detail here. The arrested state is distinguished from pausing by backtracking of the polymerase that leads to a disengagement of the 3' end of the transcript from the active site.^{6b,57–58} Evidently the energetic barrier to restarting elongation at an arrest site is greater than that which can be achieved by the polymerase alone. To reactivate elongation the enhancement of the Pol II intrinsic cleavage activity by TFIIS is required as described below.

Regulation of elongation takes advantage of the availability of the paused and arrested states. Intrinsic pause sites are found distributed fairly evenly across transcribed regions of genomic DNA with most stretches of 100 nt having a number of significant pause sites. Very strong pause sites are found in a few conserved locations such as downstream of the 3' end processing site of *Drosophila* histone genes,⁵⁹ but most sites of promoter proximal pausing, described in detail below, lack such conserved sites. Factors can either increase or decrease the dwell time at pause sites. The combination of DSIF and NELF increase pausing over intrinsic pause sites⁶⁰ and TFIIF can reduce the dwell time.⁶¹ Because the elongation rate is dictated by the pause times, DSIF and NELF reduce the elongation rate about 3 fold and TFIIF increases the elongation rate about 20 fold. For the factors studied so far new pause sites are not created nor are pause sites eliminated.

4.2 Paused Pol II on most metazoan promoters

More than two decades ago, the discovery of two distinct types of Pol II elongation modes opened the era of understanding promoter proximal pausing and productive elongation. Bentley and Groudine discovered that repression of *c-Myc* transcription is due to a block to elongation instead of initiation.⁸ Kao et al. showed that the HIV transactivator Tat up-regulates HIV transcription by relieving an elongation block without affecting initiation.⁹ A series of studies from the Lis lab showed that the promoter of *Drosophila HSP70* (and later several other genes) is occupied by transcriptionally engaged Pol II that can continue elongation upon activation or treatment with Sarkosyl in a nuclear run on experiment.^{10,62} The basics of Pol II elongation control were revealed with an in vitro system using

immobilized templates and nuclear extract.⁶³ Examination of transcription from a number of promoters demonstrated the existence of both negative transcription elongation factors (N-TEFs) that created an early block and positive transcription elongation factors (P-TEFs) that were needed to relieve the block so that the polymerases could enter productive elongation.⁶³

The phenomenon of promoter proximal pausing has been generalized in mammals and *Drosophila*,^{7,64} but is not apparent in yeast.⁶⁵ Genome-wide analyses using ChIP-chip showed that a large fraction of human and *Drosophila* genes have Pol II poised in promoter proximal positions.¹¹ The Pol II profile in mouse embryonic stem cells illustrates the three classes of transcriptional states (Figure 3).⁷ One class, exemplified by one of the olfactory receptor genes, *OLFR1080*, does not have Pol II associated at any position, indicating that the gene is shut down and generating no, or miniscule level of RNA. The second class is like the *CALM1* gene which shows promoter proximal paused Pol II with no significant Pol II signal over the gene body. This type of gene is relatively poorly expressed, but could be potentially activated like *Myc*, *HIV* and *HSP70* mentioned earlier and are, therefore, poised to respond to activation signals either from the environment or during development.^{11a,11c} The third class is constituted of several hundred highly active genes. As illustrated by *RPL3*, this class has a high signal of promoter proximal paused Pol II, a variable level of Pol II over the gene body and a broad Pol II peak at the 3' end downstream of the Poly(A) signal. Treatment of cells with the P-TEFb inhibitor flavopiridol has no effect on promoter proximal pausing, but essentially eliminates polymerases downstream of the paused Pol II as is obvious for *RPL3* (Figure 3, Pol II + Flavo). A study analyzing *Drosophila* embryos at different stages of differentiation showed that many genes could switch from one class to another during development.⁶⁶

Sequencing strand-specific nascent transcripts from human cells revealed a new widespread feature for promoters: they appear to initiate in both sense and anti-sense directions from TSSs that are about 200 bp apart.^{53,67} The anti-sense transcription is, in general, found to be lower in level^{53,67a,b} and less stable.^{67c,68} Divergent transcription can be seen in the metagene analysis in the bottom of Figure 3. On average, Pol II peaks at about +70, but an upstream hump found at about -250 is due to divergent transcription. On average, TBP and the MED1 subunit of Mediator are found between the two peaks of Pol II (Figure 3). Transcription in both directions encounters promoter proximal pausing as indicated by NELF and DSIF co-occupancy with Pol II.⁶⁸⁻⁶⁹ When P-TEFb is inhibited by flavopiridol, nascent transcripts on both directions decreased with similar kinetics for tested genes, suggesting both directions undergo P-TEFb dependent pause-release.⁶⁸ However, it is to be elucidated why processed mRNA is almost exclusively produced from the "sense" direction, although exosome degradation of "anti-sense" transcripts is noted to be involved.^{67c,68} Genome-wide sequence analyses indicated that, at locations close by the TSSs, Poly(A) signals are depleted whereas U1 snRNP sites are enriched in the sense direction, suggesting transcription in the sense direction is less likely to terminate pre-maturely.⁷⁰ It is also shown that >60% of long noncoding RNA is generated from divergent transcription in human and murine embryonic stem cells.⁷¹ Interestingly, transcription of sense and divergent promoter pairs appear to be coordinately regulated.⁷¹ Divergent transcription is also observed in yeast,

where it is shown to be a major source of unstable non-coding RNA.⁷² A recent study using both ChIP-Seq and GRO-Seq showed that divergent promoters are far less prevalent in *Drosophila*.⁵⁴ Analyses of both human and *Drosophila* data indicate that promoters with directional sequence elements (e.g. TATA box) tend to be transcribed unidirectionally and promoters containing only nondirectional motifs (e.g. CpG islands) tend to be transcribed bidirectionally.^{54,73}

4.3 Factors involved in promoter proximal pausing

4.3.1 NELF and DSIF—Promoter proximal pausing is, in part, caused by the action of two factors, the DRB sensitivity-inducing factor (DSIF) and the negative elongation factor (NELF) (Figure 4). Several recent reviews provide a detailed overview and perspective of NELF and DSIF.⁷⁴ The small molecule DRB (5,6-Dichlorobenzimidazole 1- β -D-ribofuranoside) that inhibits P-TEFb⁷⁵ and mRNA production⁷⁶ and later was found to specifically inhibit transcription elongation, was key to their discovery.⁷⁷ In vitro transcription systems suggested that initiated Pol II elongation complexes come under the influence of negative elongation factors in the presence of DRB.⁶³ DSIF was later purified and shown to be required for the effect of DRB in vitro⁷⁸. NELF was then identified as cooperating with DSIF to convey the negative effect on elongation.⁷⁹ Recombinant DSIF and affinity purified NELF together reconstituted a strong negative effect on Pol II elongation in vitro.^{47b,60,80} At the P-TEFb mediated transition into productive elongation, NELF is phosphorylated and released from the elongation complex⁸¹, whereas DSIF is phosphorylated and tracks with elongating Pol II as a positive elongation factor.⁸² Genome distributions of DSIF and NELF in MES cells⁶⁹ confirmed that NELF is only present in promoter proximal regions whereas DSIF is also present over gene bodies and at 3' end of genes, echoing that of Pol II (Figure 3).^{16a} Notably, both DSIF and NELF are present on Pol II in both sense and anti-sense directions at divergent promoters (Figure 3).⁶⁸ Knockdown of NELF reduced the level of Pol II promoter proximal pausing, although the degree of effect varies between genes.^{11c,54} Two recent studies have added to our understanding of the mechanism of NELF function during promoter proximal pausing. In *Drosophila* NELF was found to interact with DNA binding transcription factors, GAF or MIBP, and this affects the timing of loading of NELF onto the polymerase and, therefore, the location of the pause.⁸³

The degree of evolutionary conservation of DSIF and NELF are significantly different. DSIF is highly conserved as evidenced by a prokaryotic homologue, NusG, that is involved in transcription elongation.⁷⁴ The highly conserved regions of the Spt5 subunit of DSIF are involved in interactions with the polymerase that bridge the two sides of the active center cleft and help maintain stability of the elongation complex.⁵⁷ NELF seems to be a more recent evolutionary innovation. The four subunits of NELF can be found in most metazoans, but are missing from some nematodes.^{74a} NELF subunits are absent in yeast, plants, and prokaryotes. The role of NELF as a pausing factor is supported by this phylogenetic distribution because where NELF is absent, the organisms do not exhibit promoter proximal pausing of Pol II.

4.3.2 Gdown1 and GNAF—The recently characterized substoichiometric Pol II subunit, Gdown1, has been implicated in generating stably paused Pol II at promoters in vitro and in

vivo (Figure 4).^{39a} Gdown1 can be incorporated into Pol II elongation complexes in vitro and remain stably associated (resistant to 1.6M salt wash) and the resulting Pol II(G) is resistant to elongation stimulation by TFIIF (see Section 5.4) and resistant to termination by TTF2.^{39a} Acting alone Gdown1 does not have a significant effect on elongation, but it imparts a strong negative effect in the presence of nuclear extract. During fractionation of nuclear extract a Gdown1 negative accessory factor (GNAF) was uncovered that worked with Gdown1 to confer the negative elongation effect.^{39a} ChIP-Seq data showed that the genomic distribution of Gdown1 correlates with that of Pol II with strong occupancy over promoter proximal regions.^{39a} Because Pol II(G) was originally shown to be a subpopulation of total Pol II,³⁸ it remains to be clarified what other functional distinctions exist between the two forms of elongation complexes. During mitosis all Pol II molecules are terminated by TTF2,¹⁷ and the Pol II population should include Pol II(G).^{39a} This strongly suggests that there is a mechanism to control termination of Pol II(G). One possibility is regulation by post translational modification and supporting this hypothesis, Gdown1 has been demonstrated to be phosphorylated.^{39b} If such regulated termination exists during mitosis it is possible that termination of Pol II(G) during normal interphase transcription or during S phase might also be controlled.

4.3.3 TFIIS—The first Pol II transcription factor discovered was TFIIS (Figure 4).⁸⁴ The factor allowed Pol II to extend transcripts beyond arrest sites by stimulating the intrinsic transcription cleavage activity of the polymerase.⁸⁵ Factors that induce transcript cleavage activity by RNA polymerases are conserved from bacteria to man.⁸⁶ The cleavage of nascent transcripts in backtracked elongation complexes allows the 3' end of RNA to align with the active site and thereby resume elongation. The positive aspect of TFIIS function is discussed below, but in some circumstances TFIIS can manifest as a negative activity. In vitro, the inherently negative activity (transcript shortening) can overwhelm the positive effect of NTPs on elongation.^{39a} TFIIS was re-discovered during the fractionation of factors required for the negative effect of Gdown1 on transcription elongation.^{39a} Gdown1 does not directly influence TFIIS activity and in the absence of strong positive elongation factors TFIIS helps to restrict elongation to promoter proximal regions in vitro.^{39a} In vivo, *Drosophila* TFIIS has been shown to be required for efficient release of paused Pol II,⁸⁷ but when TFIIS was knocked down the short nascent transcripts associated with paused Pol II became longer providing evidence that TFIIS might be involved in restricting the movement of paused Pol II.⁸⁸

4.4 Impact of pausing on promoter function

4.4.1 Rapid response—One of the first recognized benefits to having promoter proximal paused Pol II is having a pool of Pol II elongation complexes for prompt responses to activation signals. This mechanism allows expression of genes like *c-Myc*, *HSP70* and *HIV* to be quickly amplified.^{8–10,62,89} Signals could be developmental and/or tissue specific.^{11a} One set of studies have demonstrated that genes with paused polymerases can be synchronously induced, thereby ensuring that all cells respond at the same rapid rate.⁹⁰ It has been demonstrated that in the lipopolysaccharide inducible expression programs, primary response genes, in contrast to secondary response genes, have promoter proximal paused Pol II and respond to activation signals by quickly recruiting P-TEFb and

subsequently generating mature mRNAs.^{12a} Genes that activated by estrogen commonly have promoter proximal paused Pol II and estrogen enhances the transition into productive elongation.^{12c} Comparison of different inducing reagents and cell lines showed that the resulting rate of productive elongation varies by as much as 4-fold depending on activation methods, cell lines, and specific genes, indicating further regulation after the response.^{14c}

4.4.2 Maintaining open promoters—Insight into another benefit to having promoter proximal paused Pol II came from studies in which the pausing factor NELF was reduced in cells. Knockdown of two NELF subunits, NELF-B and NELF-E, in *Drosophila* S2 cells affected the expression of a subset of genes and majority of the affected genes displayed a reduction in expression. Pol II pausing and chromatin marks indicative of active promoters were reduced or eliminated on the down regulated genes. Thus, NELF is suggested to play a positive role in keeping a permissive landscape for transcription.^{11c,91} Similar results were found in mouse embryonic stem cells after knockdown of the NELF-A subunit.⁶⁹ Therefore, one property of paused Pol II elongation complexes is to keep promoters in an active, open configuration. In this way the paused elongation complexes are seen to direct the chromatin structure and should be considered an important chromatin mark with distinct properties. One of these properties is to act as chromosomal boundaries between distant control regions and other promoters.⁹²

5. Productive elongation

Elongation complexes paused in promoter proximal positions are the targets of P-TEFb which provides the trigger to enter productive elongation. About 20 years ago, the biochemical separation of negative and positive transcription elongation control factors led to the generation of a model for elongation control⁶³ which was followed by the purification of P-TEFb^{13a} containing CDK9⁹³ and Cyclin T in *Drosophila*⁹⁴ or Cyclin T1 and T2 in humans.⁹⁵ Although P-TEFb is a CTD kinase,^{13b} it is probably phosphorylation of the Spt5 subunit of DSIF^{44b,46,82} or the RD subunit of NELF⁸¹ rather than the large subunit of Pol II that leads to the recruitment of factors needed for the rapidly elongating form of Pol II that almost unwaveringly reaches the 3' ends of genes. In yeast, where NELF is absent, the CDK9 homolog Bur1 triggers the transition by phosphorylating Spt5 as well as the CTD.⁹⁶ The P-TEFb requirement is fulfilled at the 5' end of genes such that continued P-TEFb kinase activity is not required once productive elongation has begun.^{17,75} There is a dramatic exchange of factors as promoter proximal paused Pol II makes the transition into productive elongation. Phosphorylated DSIF is retained, but NELF is lost and other factors are added (Figure 5).⁶⁹

5.1 Elongation rate

The most significant property of productive elongation complexes is the high rate of nucleotide addition. Although it is difficult to determine the elongation rate for a promoter proximal paused Pol II, numerous *in vivo* observations and *in vitro* assays suggest that promoter proximal paused Pol II is essentially static or moving less than 5 bases per minute.^{16a} Pure mammalian Pol II in the absence of other factors pauses at many positions along a dsDNA template *in vitro* and the resulting elongation rate is 25 to 50 nt/minute.⁹⁷

Initial estimates in cells suggested the in vivo elongation rate is 3–6 kb/min.⁹⁸ On immobilized templates, productive elongation complexes reached run-off at 0.5 kb within 1 min of P-TEFb activation.^{41b,99} In vivo measurements of mammalian Pol II elongation rate using various methods yield a rate of 1.3–4.3 kb/min.^{14a} One indirect method found a rate more than an order of magnitude higher on single copy integrated reporter genes.¹⁰⁰ Studies surveying elongation rate of long natural human genes showed the average rate is ~3.8 kb/min by RT-PCR¹⁰¹ and 3.1 kb/min by tiling arrays.¹⁰² A recent GRO-Seq study measuring “waves” of inducible transcription concluded similar numbers on average with variations between specific genes, cell lines, and induction methods.^{14c} As described below nucleosomes constitute a significant barrier to elongation in vitro and understanding how Pol II overcomes these blocks in vivo is needed to explain the high rates of elongation in cells.¹⁰³

5.2 Super Elongation Complex

The exact composition of the productive elongation complex, the order in which the factors are incorporated and removed, and the stability of their interactions are only partially understood, although significant progress has been made in identifying new factors (Figure 5). A series of P-TEFb containing protein assemblies, termed super elongation complexes (SECs), have been identified and characterized by several groups as reviewed.^{16b} SEC components include AF4/FMR2 family (AFF) member AFF1 and AFF4, eleven-nineteen Lys-rich leukemia (ELL) family member ELL1, ELL2 and ELL3, eleven-nineteen leukemia (ENL), and ALL1-fused gene from chromosome 9 (AF9). ELL1 was initially identified to enhance Pol II elongation by suppressing pausing¹⁰⁴ and later suggested to regulate leukemic pathogenesis via SEC.¹⁰⁵ Different forms of affinity purified SECs from stably transfected cells could stimulate transcription elongation in vitro.¹⁰⁶ SEC components were later generalized to be required for induction of many mammalian genes, as suggested by the distribution of AFF4 ChIP-Seq signals.¹⁰⁷ SEC should be considered a blanket term to describe a variety of factors involved in productive elongation rather than a defined biological complex. ENL and AF9 were recently shown to not co-exist in the same SEC.¹⁰⁸ Also, homologous proteins (ELL1 and ELL2, AFF1 and AFF4) exist in different SECs and, therefore, create the opportunity for more gene specific regulation.^{106,109} With the identification of *Drosophila* SEC, another form of ELL-containing complex that lacks P-TEFb showed enrichment on small nuclear RNA genes.¹¹⁰ AFF2 and AFF3, proteins related to AFF1 and AFF4, are also suggested to associate with P-TEFb and AF9/ENL in human cells, possibly regulating expression of different sets of genes.¹¹¹ Hinting at a more general role in gene activation, ELL3 has been shown to co-localize with other important transcription factors at enhancers in embryonic stem cells.¹¹² Much work is needed to sort out the distribution of the different SECs across genes and to determine the biochemical mechanisms of their action.

5.3 Paf1 complex

The human Paf1 protein (Pol II-associated factor) is part of the Paf1 complex (Paf1C) that additionally contains Ctr9, Cdc73, Rtf1, Leo1 and Ski8 as reviewed (Figure 5).¹¹³ The Paf1C was initially identified in yeast where it associated with Pol II to regulate transcription.¹¹⁴ It has conserved functions in regulating histone modifications, transcription

elongation and mRNA processing as reviewed.^{113a} A detailed biochemical study showed that Paf1C works with DSIF in a non-redundant manner to stimulate Pol II elongation in vitro and in vivo.^{44b} Positive effects from both factors require P-TEFb action. Paf1C has also been shown to stimulate transcription through a recombinant chromatin template with the help of p53 and histone acetyltransferase p300.¹¹⁵ The Paf1 subunit co-immunoprecipitated with the chromatin interacting ENL and AF9, which led to the proposal that Paf1C connects SEC to Pol II on chromatin.¹⁰⁸

5.4 TFIIF function

TFIIF, consisting of two subunits RAP30 and RAP74, was originally identified as a Pol II binding factor.¹¹⁶ In addition to its function in initiation both in vivo and in vitro (recently reviewed¹¹⁷), TFIIF also affects elongation significantly in vitro.^{61,118} It will bind to paused elongation complexes, but does not remain bound during elongation in vitro.^{39a,46,61} These data support the hypothesis that TFIIF interacts with paused Pol II and causes a conformational change that facilitates the re-entry into elongation mode.⁶¹ This notion is supported by the flexible and dynamic interaction between the yeast Pol II and TFIIF.¹¹⁹ A recent in vitro study indicated that TFIIF was needed to stabilize the interaction of TFIIB with Pol II during very early elongation.²⁹ Interestingly, phosphorylated TFIIF has much reduced activity in elongation stimulation but maintains its function in PIC formation, suggesting a functional switch between being an initiation factor and an elongation factor.⁴³ Although genome distribution of yeast TFIIF by high resolution ChIP-exo showed that it is mostly concentrated on promoter regions in yeast,^{65a} enrichment at downstream regions could be found in human cells.¹²⁰ The role of TFIIF in elongation in vivo could be regulated by Gdown1, which blocks TFIIF function during elongation in vitro and shows enrichment in gene bodies by ChIP-Seq.^{39a} A high quality TFIIF ChIP-Seq dataset in human cells would certainly be beneficial in resolving the role of TFIIF in elongation.

5.5 TFIIS function

In vitro transcription assays documented that Pol II elongation frequently pauses at certain sites on the template.¹²¹ This transient pausing, if not released in time, could lead to Pol II backtracking and the arrested state.¹²² TFIIS, an elongation factor that directly interacts with Pol II,¹²³ facilitates release of Pol II from arrest⁸⁷ by stimulating the RNA cleavage activity of the polymerase.⁸⁶ Overexpression of a cleavage-defective mutant causes lethality in yeast indicating that transcript cleavage is essential for cell viability.¹²⁴ Structural studies detailed that TFIIS displaces backtracked RNA in the secondary channel (pore and funnel) and stimulates the cleavage activity.¹²⁵ This process is included in a movie produced by the Cramer lab that captures many key aspects of Pol II initiation and elongation.^{6b} TFIIS was inhibited in vitro by the combination of DSIF and NELF,¹²⁶ whilst it stimulated elongation with the assistance of other cellular factors.⁸⁷ Furthermore, a nascent transcript of at least 13 nt is required for TFIIS activity in vitro.¹²⁷ Knockdown of TFIIS in *Drosophila* cells lead to a reduced efficiency of release of paused Pol II.⁸⁷ Other studies also indicated involvement of TFIIS during productive elongation. TFIIS can bind to and cooperate with the Paf1C complex in stimulating transcription elongation in vitro,¹¹⁵ and inhibition of the TFIIS-Paf1C interaction reduced expression of a set of pro-oncogenes in HeLa cells.¹²⁸

5.6 Mediator and Gdown1

Mediator's connection with Pol II elongation emerged from different directions. Firstly, it should be noted that Mediator, like SEC, may not exist or function as a single "complete" complex at all times.^{35b,129} Many individual subunits in different organisms have been studied in the past and a unified nomenclature was re-established for clarity in the field.¹³⁰ Mediator appeared to be in different conformations when bound with different activators.^{37a,129} Purifications of tagged Mediator subunits yielded varied levels of different components.¹³¹ Immunoprecipitation of the tagged SEC component AF4 detected SEC and Mediator components, hinting at a more complex and dynamic interactome during elongation.¹⁰⁶

Individual Mediator components have recently been assigned significant functions. MED26, a subunit for a subset of metazoan Mediator, interacts with TFIID and SEC components with its N terminal domain.¹³² MED26, therefore, is seen to facilitate the function of connecting initiation and elongation, possibly as part of the conformational transition of the transcription complex. MED23 interacts with the splicing machinery and regulates alternative mRNA processing.¹³³ In the search for functional targets of MED23, it was revealed that the presence or absence of MED23 oppositely regulates two sets of genes to achieve distinctive cellular outcomes.¹³⁴ A mutation in the human *MED23* gene impaired JUN and FOS in responding to an activation signal and has been suggested to be linked to intellectual disability.¹³⁵ MED12 mutations are linked with dysregulation of transcriptional response¹³⁶ and tumorigenesis.¹³⁷ A following study identified MED12 to be responsible for controlling the response to multiple cancer drugs.¹³⁸

Mediator has a sub-complex consisting of CDK8, cyclin C, MED12 and MED13 which is sometimes called the CDK8 module or the kinase module.¹³⁹ The overall function of the kinase module has not been clearly defined. Early studies suggested that it had a negative effect on initiation in vitro and acted as a repressor of gene expression, but more recently it has been shown to positively influence transcription.^{35b,129,140} Biochemical analyses have determined that the interaction of the kinase module and Pol II with Mediator are mutually exclusive.¹⁴¹ However, the genome wide distribution of the body of human Mediator (MED1) and the kinase module (MED12) are almost indistinguishable.^{37b} Uncovering the positive role of the module, CDK8 is recruited upon activation and is required for induction in human cell lines.¹⁴² CDK8 knockdown reduced Brd4 recruitment and Pol II CTD phosphorylation on Ser5 and Ser2 residues, suggesting its function in regulating elongation and this is supported by the finding that Mediator containing the CDK8 module interacts with P-TEFb.¹⁴³

Gdown1's impact on elongation is most definitively shown in vitro.^{39a} Joining and becoming tightly associated with isolated early elongation complexes, Gdown1 inhibits elongation stimulation by TFIIF and inhibits TTF2 driven termination.^{39a} Biochemical and structural studies showed that Gdown1 directly interacts with the RBP1 and RBP5 subunits of Pol II.^{39b,c} Gdown1 is brought into the elongation complex probably through its connection with Mediator during initiation.¹⁴⁴ ChIP-Seq data revealed that Gdown1 globally co-localizes with promoter proximal paused Pol II and is also present over gene bodies to

varying degrees.^{39a,b} The level of Gdown1 over gene bodies is reduced in cells treated with the P-TEFb inhibitor flavopiridol, suggesting that it can be part of productive elongation complexes, however its exact role in productive elongation remains unclear.^{39a}

5.7 Chromatin versus elongation

Just as the default chromatin structure obstructs Pol II from accessing promoters, unmodified nucleosomes present a challenge to the polymerase during elongation.^{103d} Promoter proximal paused Pol II and nucleosomes cannot co-occupy the same position, therefore, when elongation is paused, Pol II wins the competition for genome occupancy.¹⁴⁵ This dominant position is used to direct modification of surrounding nucleosomes with the H3K4ME3 histone mark.¹⁴⁶ The question of what happens when an elongating polymerase encounters a nucleosome has been examined thoroughly in vitro.^{103c,103f,147} In the absence of elongation factors Pol II can only penetrate about 50 nt into a nucleosome before it stalls.^{103d} Addition TFIIF or TFIIS alone have a slight positive effect and a little more if added together.¹⁴⁸ However, the rate of elongation through a nucleosome in vitro is more than two orders of magnitude slower than that found in vivo.¹⁴⁸ A factor that facilitates transcription through chromatin, FACT, enhances the ability of Pol II to traverse a nucleosome in vitro, but again the rate is relatively low compared to the in vivo rate.¹⁴⁹ So far, in vitro assays examining elongation of Pol II through a nucleosome have all used Pol II along with a few defined factors. It has not been determined if Pol II with all the factors needed to maintain P-TEFb-dependent, productive elongation is able to traverse a nucleosome in vitro.

Analysis of chromatin in cells demonstrates that transcribed regions are different from non-transcribed regions. Concerted modifications on different residues of the histones are found throughout the bodies of transcribed genes, but the effect of these modifications on elongation is less clear.¹⁵⁰ One possibility for achieving high elongation rates in vivo is removal of nucleosomes and highly active genes have been demonstrated to be at least partially depleted of nucleosomes. ChIP-chip and ChIP-Seq experiments demonstrate that active genes can be occupied with both Pol II and nucleosomes. However, because there are millions of cells in each analysis, the results are averages across the population and individual genes may not have both histones and Pol II. One of the most highly studied genes, *Drosophila* HSP70 displays a depletion of nucleosomes after heat shock induction.¹⁵¹ Surprisingly, this depletion also occurred in the presence of DRB presumably before the gene body was transcribed.¹⁵¹ Other results suggest that a pioneer round of transcription may be involved in modifying the histones.¹⁵² Studies in yeast have resulted in a model in which nucleosomes are removed ahead of the advancing Pol II and deposited behind,¹⁵³ which could explain little or no net loss of nucleosomes observed on some genes.¹⁵⁴ The dynamic nature of nucleosomes over transcribed regions is further evidenced by the incorporation of histone variants H2A.Z and H3.3.^{150a} H2A.Z can be replaced without removal of the entire nucleosome, but H3 requires removal of the core H3/H4 tetramer and suggests removal of the nucleosome occurs at least temporarily.¹⁵⁵

6. The 7SK snRNP

Because of the large number of human genes that are occupied with promoter proximal Pol II, P-TEFb must be fastidiously controlled and to a large part this is accomplished by reversible association with the 7SK snRNP. When active P-TEFb (modified by T-loop phosphorylation) enters the 7SK snRNP it is inhibited and kept in this inactive state until it is released.¹⁵⁶ In a rapidly dividing cell most of the P-TEFb is found in the 7SK snRNP and the rest is associated with chromatin and can be extracted with 150 mM salt.¹⁵⁷ Human 7SK snRNP has been extensively characterized¹⁵⁸ and the homologous snRNP from *Drosophila* has also been positively identified.¹⁵⁹ The amount of the 7SK snRNP in a given cell type depends on the state of growth. For example, before human T cells are activated they have a low level of P-TEFb and the 7SK snRNP. As the T cells begin to proliferate due to activation, levels of P-TEFb rise, but the new P-TEFb is found in the 7SK snRNP whose components rise concomitantly.¹⁶⁰ 7SK RNA and protein components found in the human and fly snRNP have been identified across a wide range of animal species.¹⁶¹ All organisms that have identifiable components of the 7SK snRNP also have NELF and if not already shown would at least be predicted to have promoter proximal paused polymerases.

6.1 Components of the 7SK snRNP

In addition to 7SK RNA that acts a scaffold, the 7SK snRNP is defined by its protein constituents. Two proteins, LARP7 and MEPCE, are constitutively associated while HEXIM, P-TEFb, and several hnRNPs are in dynamic equilibrium. Release of P-TEFb is accompanied by loss of HEXIM with both being replaced by hnRNPs. In this way, the 7SK snRNP is actually a mixture of complexes with varying composition (Figure 6). The Cdk9 subunit of P-TEFb must be phosphorylated on T186 of the T-loop for it to be active and for it to associate with the 7SK snRNP.^{156a,156c}

6.1.1 7SK—Early studies examining P-TEFb by glycerol gradient sedimentation analysis and immunoprecipitation demonstrated that it could be part of a larger, ribonuclease-sensitive complex and that the RNA was the prevalent 7SK snRNA that previously had no known function.¹⁵⁷ Before the role of 7SK in controlling P-TEFb was known it was physically characterized and a potential structure of the RNA was proposed that was based on sensitivity to small RNA reactive compounds and nucleases.¹⁶² Some of the modification results did not fit the original proposed structure and it was later determined that 7SK undergoes a conformational change upon loss of P-TEFb and HEXIM.¹⁶³ These results led to the understanding that the 7SK snRNP is a mixture of at least two complexes that differ in protein content and RNA structure (Figure 6).

The secondary structures adopted by 7SK RNA in its various conformations are not completely defined, but it is likely that they are intimately involved in the regulation of the release and re-sequestration of P-TEFb. Biochemical and evolutionary studies addressing the structure of 7SK RNA have been recently reviewed^{158,161,164} and although much is known, the field awaits an actual structure determination of any 7SK snRNP. What is clear is that the conformational change upon release of P-TEFb has an impact on the function of causes a conformational the RNA. HEXIM binds to 7SK RNA in vitro in the presence or absence of

P-TEFb,^{156a,165} but in cells HEXIM is present only if P-TEFb is present.^{156a,163} Release of P-TEFb from the 7SK snRNP in vitro leads to dissociation of HEXIM and the accompanying change in the structure of the RNA.¹⁶³ These results suggest that the conformation of 7SK that is able to bind HEXIM is restricted by other components of the snRNP and that an active rearrangement of the protein RNA interactions is needed to get HEXIM to associate and to recruit P-TEFb. This idea is further supported by the finding that HEXIM is in excess over the 7SK snRNP in cells^{156a} even though a significant fraction of the snRNP does not contain HEXIM or P-TEFb.¹⁶⁶ An evolutionarily conserved structural element of 7SK in which the exact 5' end is paired with a region upstream of a conserved 3' stem loop may play a role because it is the anchor for a large stem loop comprising the HEXIM binding region in the first 100 nt of the RNA.¹⁶¹

6.1.2 P-TEFb—The cyclin dependent kinase controlling the elongation phase of transcription can become part of the 7SK snRNP only if it is activated by phosphorylation (Thr186 in human CDK9) and is paired with an appropriate cyclin partner (Cyclin T1 or T2 in humans and Cyclin T in *Drosophila*).¹⁶⁷ The association occurs through an interaction of 7SK bound HEXIM and likely involves contacts with the RNA directly and perhaps with other components of the 7SK snRNP.^{165c,166,168} HEXIM proteins are dimers and evidence for the presence of two molecules of P-TEFb in the 7SK snRNP has been found both in vitro and in cells.^{156a,168b}

6.1.3 HEXIM—After P-TEFb, human HEXIM1 was the second protein to be identified in the 7SK snRNP.¹⁶⁹ It has been demonstrated to exist as a dimer when free or when bound to 7SK.^{156a} *Drosophila* has only one HEXIM gene, but a gene duplication event occurred during evolution before divergence of marsupials and placentals leading to two HEXIM proteins, HEXIM1 and HEXIM2.¹⁶¹ The most highly conserved protein HEXIM1 paradoxically lacks introns and was, therefore, likely derived from the intron containing HEXIM2 gene. In HeLa cells HEXIM1 is the predominate form, but when HEXIM1 is knocked down HEXIM2 takes its place in the 7SK snRNP.^{156a,165a,b}

Although RNA-binding causes a conformational change in HEXIM that allows it to bind to and inhibit P-TEFb, the exact region(s) and/or structure of 7SK or other RNAs contacted by HEXIM proteins have not been precisely determined. HEXIM can bind to any double stranded RNA and the majority of HEXIM1 that is found outside of the 7SK snRNP in HeLa cells is associated with small RNAs including at least one miRNA.¹⁷⁰ HEXIM1 can associate with HIV TAR RNA in vitro¹⁷¹ and has been found using SELEX to associate with some mRNAs through a short hairpin structure with a small bulge (like HIV TAR RNA).¹⁷² A GAUC containing motif in a long hairpin adopted by the 5' end of 7SK is important for this binding.¹⁷³ Studies in cells have indicated that the 3' stem loop of 7SK is also required for HEXIM1 binding, but it is not clear that this is a site of direct interaction.¹⁷⁴ A direct interaction has been demonstrated in vitro in that the RNA binding region of HEXIM1 (aa 210-220) crosslinks to U30 of 7SK.¹⁷⁵ Also a recent study indicated that phosphorylation of serine 158 in HEXIM1 by PKC leads to inhibition of RNA binding in vitro and release of P-TEFb from the 7SK snRNP in cells.¹⁷⁶ HEXIM1 has been shown to associate with the glucocorticoid¹⁷⁷ and androgen receptors¹⁷⁸ and with two proteins

involved in p53 metabolism, nucleophosmin and HDM2¹⁷⁹ and these interactions seem to be independent of RNA.

HEXIM1 got its name because it was first identified as a hexamethylene-bis-acetamide-inducible transcript in vascular smooth muscle¹⁸⁰ and then picked up a second name, cardiac lineage protein 1 (CLP-1) due to severe cardiac problems during fetal development in knockout mice.¹⁸¹ Since then HEXIM1 has been linked to a number of developmental and proliferative defects. HEXIM1 is involved in cardiac hypertrophy because release of P-TEFb from the 7SK snRNP occurs during hypertrophy.¹⁸² It is not clear if the release is a cause or a consequence of hypertrophy, however, overexpression of HEXIM1 has been demonstrated to prevent hypertrophy in mice caused by hypoxia-induced hypertension¹⁸³ and HEXIM1 haplodeficiency caused increased susceptibility to cardiac stress but improved injury induced muscle regeneration.¹⁸⁴ Underexpression of HEXIM1 is associated with metastatic breast cancers and subsequent induction leads to an inhibition of metastasis.¹⁸⁵ The effects of HEXIM proteins are widespread in the animal kingdom. Global knockdown of *Drosophila* HEXIM is embryonic lethal and tissue specific knockdown severely impacts the development of all organs examined.¹⁵⁹

6.1.4 MEPCE—Although it was known for many years that 7SK RNA was modified by methylation of the 5' gamma phosphate,¹⁸⁶ the identity of the enzyme responsible for the unusual 5' cap structure remained a mystery until recently. A previously uncharacterized methyltransferase was among many proteins chosen for proteomic analysis and after finding it associated with P-TEFb and HEXIM proteins, it was determined that its substrate was 7SK and it was named the methyl phosphate capping enzyme, MEPCE.¹⁸⁷ The gamma phosphate methyl cap on 7SK evidently protects the 5' end from degradation because knockdown of MEPCE in cells leads to destabilization of 7SK RNA and an increase in P-TEFb not associated with the 7SK snRNP.^{187–188} MEPCE has been found associated with U6 RNA and its association with the 7SK snRNP involves contacts with LARP7 in addition to the RNA.¹⁸⁸ Although LARP7 interacts with MEPCE and inhibits methyltransferase activity, it is not known what function this serves since there has been no evidence that the methylation state of 7SK changes once it is methylated.¹⁸⁸ The function of MEPCE is well conserved as it is found associated with 7SK in *Drosophila*,¹⁸⁹ however, it is also known as the Bicoid interacting protein (Bin3) and has been found to bind to and regulate the translation of specific mRNAs.¹⁸⁹

6.1.5 LARP7—The La related protein LARP7 was also identified in the proteomic analysis of proteins associated with MEPCE¹⁸⁷ and soon after was shown to be a *bone fide* component of the 7SK snRNP.^{166,190} Like MEPCE, LARP7 has a stabilizing effect on 7SK RNA as evidenced by reduction of 7SK caused by LARP7 knockdown.^{166,190a,191} Like La and other members of the LARP family, LARP7 has a La motif that is used to interact with the U stretch at the 3' end of the RNA.¹⁹² LARP7 is almost exclusively associated with 7SK and remains associated after release of P-TEFb and HEXIM.^{159,166} Because of this, LARP7 is a prime candidate for the protein that could hold 7SK in a conformation that would be refractory to binding HEXIM and re-sequestering P-TEFb.¹⁶³ Mutations in LARP7 have

been linked to cancer and developmental defects and this could be due to misregulation of P-TEFb.^{191,193}

6.1.6 hnRNPs and other proteins—When P-TEFb and HEXIM are released from the 7SK snRNP the overall size of the complex does not change significantly due to the association of hnRNP proteins.^{164,166,194} hnRNP A1, A2/B1, R and Q are all found to increase in the 7SK snRNP that does not contain P-TEFb.^{166,194b} These hnRNPs are not specific for the 7SK snRNP and it is not clear if they directly participate in the regulation of release or re-sequestration of P-TEFb and HEXIM.^{158,164} Because HEXIM is released from immunoprecipitated 7SK snRNP upon extraction of P-TEFb by HIV Tat or P-TEFb binding domain of Brd4 in the absence of excess hnRNPs, they are not required in the release process or the subsequent conformational change in 7SK RNA.¹⁶³ Whether they are just place holders or play an active role in 7SK snRNP metabolism will require further study. Another protein, RNA helicase A (RHA), is also associated with the 7SK snRNP after P-TEFb is released and could be involved in the rearrangement of 7SK RNA that allows HEXIM to bind.^{194b}

6.2 Release of P-TEFb from the 7SK snRNP

One of the major mysteries surrounding the function of the 7SK snRNP is how the release of P-TEFb is regulated. Two proteins with strong P-TEFb interaction domains, HIV Tat and Brd4, have been demonstrated to release P-TEFb directly in vitro and when overexpressed in cells.^{163,171} The structure of HIV Tat bound to P-TEFb indicates that the primary interaction with P-TEFb is through the Cyclin T1 domain, although some contact is made with the T-loop on the CDK9 subunit.¹⁹⁵ Brd4 likely interacts with a similar surface because Tat and Brd4 compete for binding to P-TEFb.¹⁹⁶ While the interaction between HEXIM1 and P-TEFb has not been determined, there is likely some overlap with the Tat and Brd4 binding site. The mechanism whereby HEXIM1 is displaced by the extractor protein is not understood, but because it can occur in vitro in the absence of phosphate, methyl, or acetyl donors it must not require covalent modification.¹⁶³ Because HEXIM is ejected after P-TEFb is extracted in vitro, it is possible that conformational strain in the RNA induced by other proteins in the 7SK snRNP is part of the driving force for extraction.

To the extent that it has been investigated, anything that impairs the elongation phase of transcription by Pol II causes the global release of P-TEFb from the 7SK snRNP. This release could be either an increase in actual release or a decrease in re-sequestration. Treatment of cells with UV light, the DNA intercalation agent actinomycin D, topoisomerase inhibitors, P-TEFb inhibitors, Pol II inhibitors like α -amanitin.^{16a} Besides inhibiting elongation it is not clear what all the types of treatment have in common or what the actual signal or the mechanism for the release is. This is complicated by the fact that physical blocks to elongation (UV damage, actinomycin D) or P-TEFb inhibitors that only stop the transition into productive elongation, elicit the same response. It has been found that such treatments lead to a rapid increase in productive elongation from the HIV-LTR that is loaded with promoter proximal paused Pol II, presumably due to the excess free P-TEFb.¹⁹⁷ Treatment of cells with histone deacetylase inhibitors (HDACi) such as TSA or SAHA also leads to release of P-TEFb.¹⁹⁸ Acetylated histones are considered as marks for active

transcription units and one would think that inhibiting deacetylation would increase transcription. However, treatment with HDACi has been shown to repress transcription of a number of genes¹⁹⁹ and this repression must be what leads to release of P-TEFb. Because T-loop phosphorylation is required for P-TEFb to be retained in the 7SK snRNP, treatment with phosphatases in vitro also causes release of P-TEFb.^{156b}

6.3 Dynamics and subcellular localization of the 7SK snRNP

The biochemical and cellular properties of the 7SK snRNP set it apart from other snRNPs and these need to be addressed in building models for the function of the complex. The 7SK snRNP is readily released from non-ionic detergent treated nuclei unlike all other snRNPs that are tightly retained.²⁰⁰ This suggests that it is not tightly or stably associated with chromatin. This release occurs even at very low salt (10 mM).²⁰⁰ However, P-TEFb that is not in the snRNP is retained by nuclei until the salt is raised to over 150 mM.²⁰⁰ Because this includes P-TEFb globally released by agents that block elongation, it suggests that P-TEFb not in the 7SK snRNP is associated with chromatin. Because Brd4 is bound to chromatin and can associate with P-TEFb it is possible that the released P-TEFb is bound to Brd4. Brd4 is more tightly bound to chromatin (released above 300 mM salt) than P-TEFb and this means that it is not possible to extract the Brd4•P-TEFb complex from nuclei intact. Interpretation of studies that concluded that most P-TEFb released from the 7SK snRNP is associated with Brd4 should be tempered with the fact that these studies^{196,201} required dialysis of high salt extracts to observe the interaction between Brd4 and P-TEFb. Because Brd4, but not P-TEFb is associated with mitotic chromosomes, the interaction must be regulated in some way.

Both P-TEFb and components of the 7SK snRNP are located primarily in the non-nucleolar nuclear compartment. How P-TEFb is recruited to activate genes is not well understood. P-TEFb subunits can be found over the bodies of highly expressed genes suggesting that they are associated with elongation complexes. One group has found evidence for recruitment of the 7SK snRNP upstream of the HIV TSS.²⁰² This is not contradictory to the general biochemical properties described above because the crosslinking agent used in ChIP could stabilize a weak or transient interaction. One very interesting study examined the dynamics of 7SK localization within the nucleus when an array of genes was activated.²⁰³ Although no increase of 7SK was found upon induction of the genes, the local concentration of 7SK around the array locus increased upon shut down of the genes. This supports the idea that re-sequestration of P-TEFb is an important regulated step and suggests that there is a mechanism to recruit the 7SK snRNP, perhaps to the 3' end of genes to ensure that free P-TEFb is not made globally available.

7. Coupling of transcription and mRNA processing

7.1 CTD phosphorylation during elongation and mRNA processing

During transcription, the C-terminal domain (CTD) of Rbp1, the catalytic and largest subunit of Pol II, undergoes dynamic modifications. The CTD consists of tandem heptad repeats that fall into the consensus of YSPTSPS. There are 52 repeats in human, 26 repeats in budding yeast and intermediate numbered in other eukaryotes. The relatively unstructured

CTD lies underneath the RNA exit channel²⁰⁴ and undergoes reversible phosphorylation events at multiple positions during a round of transcription as reviewed.²⁰⁵ The timing of specific phosphorylation and dephosphorylation events crucially regulate transcription and RNA processing events (Figure 7).^{15b,c}

The CTD is a substrate for a large number of kinases *in vitro*, although the physiological significance of these findings is not uniformly clear.²⁰⁶ Five of the 7 amino acids in the heptad repeats can be phosphorylated. Tyr1 can be phosphorylated²⁰⁷ by human Abl1 and Abl2²⁰⁸ and in yeast Tyr1 phosphorylation is tied to termination.^{207b} Ser2 phosphorylation is found over the body of transcribed genes and downstream of the Poly(A) addition site in yeast^{65b} and mammals⁶⁹ and was originally thought to be carried out by P-TEFb.^{13b,15a} Recent results in *Drosophila* and human point to CDK12/Cyclin K as the true Ser2 kinase in cells.²⁰⁹ The fact that the P-TEFb inhibitor flavopiridol inhibits Ser2 phosphorylation in cells could be because it inhibits productive elongation complex formation and, therefore, eliminates the potential substrate for CDK12. Thr4 is phosphorylated by the kinase Plk3 in mammalian cell lines and Thr4 phosphorylation is found predominately over the 3' ends of highly expressed genes.^{206,210} Flavopiridol also inhibits Thr4 phosphorylation,^{15a,210} but as described for Ser2 phosphorylation that could be due to loss of the productive elongation and the substrate. Ser5 and Ser7 are most likely phosphorylated by the CDK7 subunit of TFIIF.²¹¹ Additionally, the Mediator subunit CDK8 has been shown to phosphorylate Ser2 and Ser5 *in vitro* and knock-down of CDK8 reduced the levels of Ser2 and Ser5 phosphorylation *in vivo*.^{143,212}

The CTD needs to be dephosphorylated to allow Pol II to enter next round of transcription. Several phosphatases regulate CTD dephosphorylation.²¹³ Yeast Fcp1 (TFIIF associating CTD phosphatase 1) has been shown to be able to dephosphorylate the CTD²¹⁴ and, therefore, allows recycling Pol II.²¹⁵ Human Fcp1 is capable of dephosphorylating both Ser2-P and Ser5-P *in vitro*²¹⁶ with the preference towards Ser2-P.²¹⁷ Another phosphatase Ssu72 was initially identified as a factor functionally related to TFIIB²¹⁸ and later shown to dephosphorylate the Ser5-P²¹⁹ and Ser7-P²²⁰. A set of small CTD phosphatases (SCP), SCP1, SCP2, and SCP3, that contain domains homologous to Fcp1 have been identified in human²²¹ and shown to preferentially dephosphorylate Ser5-P.^{217a,222} Human RPAP2 and its yeast homolog Rtr1 have been shown to dephosphorylate Ser5-P.^{211c,223} However, structures of RPAP2 and Rtr1 did not identify catalytic sites,²²⁴ implying their function could be regulated or indirect.

7.2 Capping

The first step in mRNA processing is 5' capping,²²⁵ which takes place during the initial stages of elongation as soon as the RNA emerges from the body of the polymerase (Figure 7).²²⁶ The capping enzyme hydrolyzes the 5' triphosphate to a diphosphate and adds a GMP to the first nucleotide via an unusual 5'-5' triphosphate link. Then, a methyltransferase methylates the N7 position of the transferred GMP generating the m⁷G cap. The cap protects the mRNA from degradation by 5'-3' exonucleases and is required for efficient initiation of translation.²²⁷ The nuclear cap binding complex (CBC) recognizes the m⁷G and promotes

co-transcriptional processing by interaction with components of the splicing and polyadenylation machineries.²²⁸

Yeast and mammalian capping enzymes have been studied and it is clear that there are significant differences between the two systems. The three enzymatic activities required (phosphatase, guanylyltransferase, and methyltransferase) are each catalyzed by separate proteins in yeast, whereas in mammals the phosphatase and guanylyltransferase activities are found in one protein, called the capping enzyme.²²⁹ Structural studies have indicated that there are also differences in the interaction of the yeast and mammalian capping enzyme with phosphorylated CTD peptides.²³⁰ CTD phosphorylation plays a major role in recruiting and activating the yeast guanylyltransferase²³¹ and in mammals Ser5-P stimulates the formation of the covalent GMP intermediate on the guanylyltransferase.²³² In vitro isolated elongation complexes stimulated human capping enzyme function 100,000 fold over free RNA, but removal of the CTD and, therefore, the effects of CTD phosphorylation only reduced capping activity 3-fold.^{226b} Evidently, the main interaction of the human capping enzyme is with the body of Pol II. The yeast capping enzyme has also been shown to interact with the body of Pol II in addition to the phosphorylated CTD.²³³ These results suggest there has been an evolutionary shift in the relative importance of the interaction with the CTD versus the body of Pol II.

Capping may be functionally linked with regulation of transcription elongation. Elongation factor Spt5, a subunit of DSIF stimulates capping several fold in vitro.²³⁴ The capping enzyme has also been shown to reverse the negative effect of NELF on elongation by Pol II in the presence of DSIF suggesting that the capping enzyme competes with NELF for binding to Pol II.²³⁵ Additionally, in separate studies, Myc is shown to regulate Pol II pause release⁶⁹ and mRNA cap methylation,²³⁶ implying another possible angle of regulation. A Rai1/Dxo1 mediated quality control mechanism for defectively capped mRNA has recently been characterized in yeast and human.²³⁷ The incompletely capped mRNA showed splicing and polyadenylation defects in human cells.^{237a}

7.3 Splicing

Most mammalian genes contain multiple introns and splicing is critically regulated to determine the coding sequence in the mRNA. The spliceosome is composed of five snRNPs (U1, U2, U4, U5 and U6) and associated proteins exceeding a total of 100 involved factors.²³⁸ It is not surprising that the CTD plays an important role (Figure 7). Phosphorylated CTD was found associated with splicing machinery and stimulates spliceosome assembly.²³⁹ In living cells, the CTD is required for the recruitment of splicing factor SC35 to Pol II elongation complexes.²⁴⁰ Through the years, the CTD has been shown to directly interact with several splicing factors and recently it was suggested to promote the initial assembly of spliceosomes.²⁴¹ Although the interplay between transcription elongation, CTD phosphorylation and splicing is challenging to dissect due to the sheer size of the complexes and the dynamic nature of the process, it is likely that Ser2-P is critical for efficient splicing. During induction of a number of mammalian genes one study noted a correlation of P-TEFb dependent Ser2-P and increased levels of splicing.^{12a} Additionally, splicing factors have stimulatory effects on elongation.²⁴² Alternative splicing is affected by

the rate of elongation²⁴³ as well as a novel protein complex, DBIRD that binds to Pol II.²⁴⁴ Further confirming the link between elongation rate and splicing, a recent high resolution PRO-Seq survey showed Pol II pausing at intron-exon junctions in general, with less used exons exhibiting less significant pausing.⁵⁵ Extended dwell time at intron-exon junctions could benefit recognition of splicing sites and improve accuracy.

7.4 3' end processing—The 3' ends of mRNAs are generated by a co-transcriptional processing event. As the Pol II elongation complex passes the Poly(A) addition signal of a gene the transcript is cleaved and a stretch of several hundred A's (in human) is added to the newly generated 3' end of the transcript (Figure 7). With the assistance of Poly(A) binding proteins, this feature protects the mRNA from degradation and promotes translation after mRNA export.²⁴⁵ The cleavage and polyadenylation machinery was initially thought to consist of about 14 proteins,^{245b} but recent mass-spec analysis identified around 85 proteins associated with the complex.²⁴⁶ Histone genes utilize a 3' end stem loop and a downstream sequence element to recruit a cleavage complex and recently a structure of this unique complex was solved.²⁴⁷ The cleavage is very fast and does not seem to be obligatorily coupled to elongation, but strong pause sites were found downstream of a number of *Drosophila* histone genes that were spaced exactly far enough downstream to allow efficient cleavage in vitro.⁵⁹ Polyadenylation and histone 3' end formation machineries have several factors in common, including the endonuclease CPSF73.²⁴⁸

Cleavage and polyadenylation occurs co-transcriptionally with the complex processing machinery intimately linked to the elongation complex.^{15b,249} It has been well documented that Poly(A) signals require Ser2 phosphorylated CTD for efficient processing.^{15a} One of the initial findings on integrating mRNA processing with transcription was that the CTD associates with Poly(A) factors.²⁵⁰ Pcf11, in particular, contains a CTD interacting domain that specifically recognizes the CTD with Ser2-P²⁵¹. The Ser2-P is also required for recruitment of Poly(A) factors and efficient 3' end processing²⁵². Surprisingly, some Poly(A) factors are found in promoter regions,²⁵³ supporting the model that CPSF is recruited by TFIID.²⁵⁴ These studies support the view that the Poly(A) complex is built in a stepwise manner during elongation. It could also be that detecting Poly(A) factors at both ends resulted from gene looping, where termination and re-initiation may take place in close proximity.²⁵⁵ Although the formation of the Poly(A) complex could start as early as initiation, productive elongation and the associated factors are required for efficient polyadenylation. Paf1C has been shown to be involved in 3' end processing,²⁵⁶ and responds to the activator VP16 to activate 3' end processing.²⁵⁷ Additionally, the SEC component ELL2 regulates Poly(A) factor CstF-64 and together they affect the activation of alternative Poly(A) signals.²⁵⁸

8. Termination

Termination is rigorously defined as the disruption of the ternary complex between the polymerase, template and RNA. For most eukaryotic genes this occurs either close to the TSS by transiently promoter proximal paused Pol II (premature termination) or downstream of the signal for 3' end processing (Figure 8). Human Pol II was designed to maintain stable elongation complexes during transcription of genes over 2 million base pairs long. This

means being resistant to termination for over 8 hr, because unlike DNA polymerases, Pol II is an obligatorily processive enzyme. Although the act of entering productive elongation is sometimes referred to as “increasing processivity” of the Pol II, Pol II is always processive (never distributive). This means that an elongation complex cannot be formed on dsDNA by addition of an RNA and polymerase. Because termination is essential to recycle Pol II and is an irreversible process, it must be carefully controlled.

8.1 Mitotic repression of transcription elongation by TTF2

TTF2 is an ATP-dependent factor that is capable of terminating Pol II and Pol I.²⁵⁹ Mutations in the gene encoding *Drosophila* TTF2, *lodestar*, lead to chromosome segregation defects in early embryos.²⁶⁰ As a member of the Swi/Snf family,²⁶¹ TTF2 translocates along dsDNA²⁶² and when it encounters the upstream edge of the polymerase it is thought to push it off of the hybrid and, thereby, cause termination. It will terminate Pol II stalled over thymine dimers²⁵⁹ and is similar in sequence and function to the *E. coli* mfd protein that terminates RNA polymerase stalled over damaged DNA to facilitate transcription coupled repair.²⁶³ TTF2 is predominately cytoplasmic during interphase, but rushes into the nucleus at the onset of mitosis and terminates essentially all polymerases, thereby, allowing efficient compaction of the chromosomes during metaphase.¹⁷ After knockdown of TTF2, mitosis will still occur normally in most cells even though Pol II remains engaged on the incompletely condensed chromosomes, but 3–10% of the cells experience problems that lead to chromosome segregation defects.¹⁷ Some TTF2 is present in the nucleus during interphase and it has been demonstrated to affect termination of promoter proximal paused Pol II.²⁶⁴ It is not known if TTF2 is involved in termination downstream of 3' end of genes.

8.2 Early termination during transcription

Transcription termination can occur prematurely before Pol II reaches the 3' ends of genes. Early in vitro studies found that a large fraction of polymerases terminate near the promoter in the absence of P-TEFb and this event was termed abortive elongation.⁶³ Termination of promoter proximal polymerases was demonstrated to be caused by TTF2 in vitro.¹⁷ TTF2 and the exonuclease Xrn2 have both been implicated in termination of promoter proximal paused Pol II in cells because depletion of the factors shifted Pol II ChIP-Seq signals to distal regions.²⁶⁴ The discovery that Pol II(G) is resistant to termination by TTF2 and that a fraction of promoter proximal paused Pol II is associated with Gdown1 indicates that the termination potential of all paused polymerases is unequal and could be regulated.^{39a} Genes with higher Gdown1/Pol II ratio could be more resistant to promoter inactivation by nucleosome assembly.

Yeast Pol II does not exhibit extensive promoter proximal pausing, but as is found in metazoans short transcripts are subject to termination. Mutations in exosome genes resulted in discovery of a class of short (200–600 nt) cryptic unstable transcripts (CUTs) generated by the TRAMP complex, which contains the non-canonical Poly(A) polymerase Trf4.²⁶⁵ Transcription for these non-coding RNA is terminated by the Nrd1/Nab3/Sen1 complex.²⁶⁶ Nrd1p and Nab3p recognize short sequence motifs at 5' ends of genes and trigger termination by interacting with CTD with Ser5-P (Figure 8).^{266b,267} The helicase Sen1 is capable of terminating Pol II in vitro.²⁶⁸ Bidirectional promoters are found to be the major

source of these pervasive short transcripts, suggesting this is a subset of the standard Pol II elongation complexes that undergoes an early termination pathway.⁷² A similar early termination complex has not been found in higher eukaryotes, however, a conserved component, Sen1/senataxin, has been shown to be involved in termination at the 3' ends of human genes.²⁶⁹

8.3 Poly(A) driven termination

When Pol II elongation proceeds beyond the 3' end of genes, termination releases the polymerase from the template for recycling as well as avoiding interference with transcription of downstream genes. It has been well documented that termination downstream of the mature 3' end requires a functional Poly(A) signal and involves the Poly(A) complex.²⁰ However, the mechanism that releases Pol II from the template is somewhat murky. A “torpedo” model and an allosteric model have been proposed. The “torpedo” model suggests the uncapped 5' monophosphate on the nascent transcript created by the cleavage reaction in 3' end processing initiates rapid degradation of the RNA, which promotes termination (Figure 8).²⁷⁰ This is supported by the function of 5'-3' exonuclease (Xrn2 in mammals or Rat1 in yeast) in Pol II termination.²⁷¹ Another model, termed the allosteric model, suggests the Poly(A) signal and the factors recruited cause a conformational change to the elongation complex to become termination competent. This model is supported by the observations that many Poly(A) factors are recruited at 3' end of genes when the CTD has high level of Ser2-P and both the factors and the Ser2-P promote termination.^{252a,252d,253} The accumulation of Pol II ChIP-Seq signals downstream of the Poly(A) signal also provides evidence of a change in the properties of the elongation complex.⁶⁹ Pcf11, in particular, has been demonstrated to dismantle the elongation complex by bridging the CTD and the nascent transcript, and depletion of Pcf11 causes transcriptional read-through.^{252c,272} The two models are not mutually exclusive, as the 5'-3' exonuclease Rat1 was shown to play a role in recruiting Poly(A) factors and in degradation of the nascent transcript attached to Pol II.²⁷³ It has been suggested that the terminating Pol II is recycled for the next round of transcription via looping from the termination region back to the promoter.^{255,274}

9. Transcription factories

The existence of transcription factories is supported by almost 200 published studies as described in recent reviews.²⁷⁵ The factories are defined by most as having multiple polymerases and transcribing multiple genes simultaneously. This model supports the observations that link initiation with 3' end processing and termination as it benefits Pol II recycling.^{255,274b} The original assays performed 20 years ago were based on a fluorescence imaging technique that detects nascent transcripts. In those studies 300–500 discrete foci for Pol II transcription were observed in the nuclei of human cells.²⁷⁶ The number of foci detected depends on the cell type, transcriptional state and can vary from 100 to more than 2000.²⁷⁷ By definition, but not rigorously determined, each factory has two or more Pol II molecules with calculations suggesting the number may be between 4 and 30.^{275c,277} It is now assumed that Pol II in the factories or in individual transcription complexes is more or less fixed in position and that DNA is reeled through. This makes sense because it is hard to

imagine how Pol II with all the associated elongation factors, nascent RNP and the coupled RNA processing machinery could track down the DNA with the concomitant rotation around the template every 10 bp. The structure upon which transcription factories is built is not known, but the Pol II is not easily released from the nuclear scaffold.

Because factors associated with transcription complexes are dynamic, the exact protein, RNA and DNA content of the factories is expected to be somewhat variable. Determination of the constant components requires isolation of the Pol II transcription factories and this is very difficult because the large complexes are tethered to an underlying nuclear structure. Digestion with DNase will solubilize 90% of chromatin while leaving the factories in the insoluble fraction.²⁷⁸ Recent progress was made in isolation by employing caspases. These proteases that are normally used during apoptosis evidently digest protein(s) used in this tethering. Proteomic analysis of complexes thus isolated yielded subunits of Pol II, dozens of hnRNPs, histones, a number of transcription factors, and potential nucleoskeleton proteins like actin and lamins.²⁷⁹ Other proteins that might be expected to be present such as components of Mediator or the SEC were not found, but the depth of the detection was not great as evidenced by the lack of most of the Pol II subunits.

The transcriptional status of Pol II in transcription factories has not been explicitly characterized. Most of the work on transcription factories took place before the prevalence of promoter proximal paused Pol II was widely appreciated.^{16a} Also, much of the Pol II is stably paused due to the presence of Gdown1,^{39a} suggesting a possible role of regulating factory formation. The amphipathic glycoside, saponin is a natural detergent found in plants that will extract free, unphosphorylated Pol II, but not phosphorylated Pol II which is engaged in transcription from nuclei.²⁸⁰ This suggests that active transcription is required for occupancy of factories. Moreover, treatment of cells with the Pol II inhibitor α -amanitin or the P-TEFb inhibitor DRB did not reduce the Pol II loci.²⁸⁰ Because neither treatment would eliminate promoter proximal paused Pol II, these results are consistent with transcription factories containing paused polymerases. Two recent studies confirmed this idea. Examination of the urokinase-type plasminogen activator gene before transcriptional activation uncovered a "poised factory" containing Pol II with Ser5 phosphorylation.²⁸¹ Also the dynamics of long inducible human genes indicates that gene loops on a transcription factory started with a promoter proximal paused Pol II and enlarge as Pol II advances.²⁸² Small factories containing Mediator and other factors could be built around individual promoter proximal paused polymerase complexes. These could progress in a P-TEFb-dependent manner into productive factories with a concomitant increase in size with the incorporation of productive elongation factors and the RNA processing machineries. Promoter proximal paused Pol II(G) may have a lifetime great enough to act as the seed for the formation of larger factories.

10. Key remaining questions

A tremendous amount of progress has been made in understanding Pol II elongation control, but there are still many unanswered questions. Some concern promoter proximal pausing. How does Gdown1 become associated with paused Pol II and how does this influence productive elongation? What are the roles of Mediator and TFIIS in maintaining the paused

state and what other factors are involved? The transition into productive elongation is incompletely understood. How does P-TEFb get in and out of the 7SK snRNP and how is its activity directed to specific genes? What are the factors responsible for the transition into and the maintenance of productive elongation and how are they recruited? Elongation control still needs to be integrated into other nuclear processes. What are the mechanisms of functional coupling between transcription and RNA processing? How are the inhibitory effects on nucleosomes on Pol II elongation overcome in cells? Finally, issues remain to be resolved concerning the termination process. What role do the different mechanisms play in termination of promoter proximal paused Pol II, divergent versus sense complexes, and at 3' ends of genes? How is Pol II termination during mitosis accomplished given that TTF2 is responsible and it is inhibited by Gdown1? There is no doubt that these and other important questions concerning Pol II elongation control will keep many of us occupied into the future. We expect that there are still big surprises in store!

Acknowledgments

We thank Tiandao Li for bioinformatic analyses shown. We are indebted to the reviewers who provided a tremendous amount of very useful information. This work was supported by National Institutes of Health Grant GM35500 to DHP and American Heart Association Fellowship 12POST12040106 to JG.

References

1. Sharp PA. *Cell*. 2009; 136:577. [PubMed: 19239877]
2. Loven J, Orlando DA, Sigova AA, Lin CY, Rahl PB, Burge CB, Levens DL, Lee TI, Young RA. *Cell*. 2012; 151:476. [PubMed: 23101621]
3. Young RA. *Cell*. 2011; 144:940. [PubMed: 21414485]
4. Vannini A, Cramer P. *Mol Cell*. 2012; 45:439. [PubMed: 22365827]
5. Lariviere L, Seizl M, Cramer P. *Curr Opin Cell Biol*. 2012
6. (a) Treutlein B, Muschielok A, Andrecka J, Jawhari A, Buchen C, Kostrewa D, Hog F, Cramer P, Michaelis J. *Mol Cell*. 2012; 46:136. [PubMed: 22424775] (b) Cheung AC, Cramer P. *Cell*. 2012; 149:1431. [PubMed: 22726432]
7. Price DH. *Mol Cell*. 2008; 30:7. [PubMed: 18406322]
8. Bentley DL, Groudine M. *Nature*. 1986; 321:702. [PubMed: 3520340]
9. Kao SY, Calman AF, Luciw PA, Peterlin BM. *Nature*. 1987; 330:489. [PubMed: 2825027]
10. Rougvie AE, Lis JT. *Cell*. 1988; 54:795. [PubMed: 3136931]
11. (a) Zeitlinger J, Stark A, Kellis M, Hong JW, Nechaev S, Adelman K, Levine M, Young RA. *Nat Genet*. 2007; 39:1512. [PubMed: 17994019] (b) Guenther MG, Levine SS, Boyer LA, Jaenisch R, Young RA. *Cell*. 2007; 130:77. [PubMed: 17632057] (c) Muse GW, Gilchrist DA, Nechaev S, Shah R, Parker JS, Grissom SF, Zeitlinger J, Adelman K. *Nat Genet*. 2007; 39:1507. [PubMed: 17994021]
12. (a) Hargreaves DC, Horng T, Medzhitov R. *Cell*. 2009; 138:129. [PubMed: 19596240] (b) Lin CY, Loven J, Rahl PB, Paranal RM, Burge CB, Bradner JE, Lee TI, Young RA. *Cell*. 2012; 151:56. [PubMed: 23021215] (c) Hah N, Danko CG, Core L, Waterfall JJ, Siepel A, Lis JT, Kraus WL. *Cell*. 2011
13. (a) Marshall NF, Price DH. *J Biol Chem*. 1995; 270:12335. [PubMed: 7759473] (b) Marshall NF, Peng J, Xie Z, Price DH. *J Biol Chem*. 1996; 271:27176. [PubMed: 8900211]
14. (a) Ardehali MB, Lis JT. *Nat Struct Mol Biol*. 2009; 16:1123. [PubMed: 19888309] (b) Palangat M, Larson DR. *Biochimica et biophysica acta*. 2012; 1819:667. [PubMed: 22480952] (c) Danko CG, Hah N, Luo X, Martins AL, Core L, Lis JT, Siepel A, Kraus WL. *Mol Cell*. 2013(d) Henriques T, Adelman K. *Mol Cell*. 2013; 50:159. [PubMed: 23622514]

15. (a) Hsin JP, Manley JL. *Genes Dev.* 2012; 26:2119. [PubMed: 23028141] (b) Moore MJ, Proudfoot NJ. *Cell.* 2009; 136:688. [PubMed: 19239889] (c) Perales R, Bentley D. *Mol Cell.* 2009; 36:178. [PubMed: 19854129]
16. (a) Zhou Q, Li T, Price DH. *Annu Rev Biochem.* 2012(b) Luo Z, Lin C, Shilatifard A. *Nature reviews Molecular cell biology.* 2012; 13:543.
17. Jiang Y, Liu M, Spencer CA, Price DH. *Mol Cell.* 2004; 14:375. [PubMed: 15125840]
18. Rondon AG, Mischo HE, Kawauchi J, Proudfoot NJ. *Mol Cell.* 2009; 36:88. [PubMed: 19818712]
19. Proudfoot NJ, Furger A, Dye MJ. *Cell.* 2002; 108:501. [PubMed: 11909521]
20. Richard P, Manley JL. *Genes Dev.* 2009; 23:1247. [PubMed: 19487567]
21. Kaplan N, Moore IK, Fondufe-Mittendorf Y, Gossett AJ, Tillo D, Field Y, LeProust EM, Hughes TR, Lieb JD, Widom J, Segal E. *Nature.* 2009; 458:362. [PubMed: 19092803]
22. (a) Struhl K, Segal E. *Nat Struct Mol Biol.* 2013; 20:267. [PubMed: 23463311] (b) Guertin MJ, Lis JT. *Curr Opin Genet Dev.* 2012
23. Kornberg RD. *Proc Natl Acad Sci U S A.* 2007; 104:12955. [PubMed: 17670940]
24. (a) Thomas MC, Chiang CM. *Crit Rev Biochem Mol Biol.* 2006; 41:105. [PubMed: 16858867] (b) Smale ST, Kadonaga JT. *Annu Rev Biochem.* 2003; 72:449. [PubMed: 12651739]
25. Juven-Gershon T, Kadonaga JT. *Dev Biol.* 2010; 339:225. [PubMed: 19682982]
26. Deaton AM, Bird A. *Genes Dev.* 2011; 25:1010. [PubMed: 21576262]
27. (a) Hoiby T, Zhou H, Mitsiou DJ, Stunnenberg HG. *Biochimica et biophysica acta.* 2007; 1769:429. [PubMed: 17560669] (b) Murakami K, Calero G, Brown CR, Liu X, Davis RE, Boeger H, Kornberg RD. *J Biol Chem.* 2013; 288:6325. [PubMed: 23303183]
28. (a) Buratowski S, Hahn S, Guarente L, Sharp PA. *Cell.* 1989; 56:549. [PubMed: 2917366] (b) Kostrewa D, Zeller ME, Armache KJ, Seizl M, Leike K, Thomm M, Cramer P. *Nature.* 2009; 462:323. [PubMed: 19820686] (c) Liu X, Bushnell DA, Wang D, Calero G, Kornberg RD. *Science.* 2010; 327:206. [PubMed: 19965383] (d) Sainsbury S, Niesser J, Cramer P. *Nature.* 2012
29. Cabart P, Ujvari A, Pal M, Luse DS. *Proc Natl Acad Sci U S A.* 2011; 108:15786. [PubMed: 21896726]
30. Yan Q, Moreland RJ, Conaway JW, Conaway RC. *J Biol Chem.* 1999; 274:35668. [PubMed: 10585446]
31. Kim TK, Ebright RH, Reinberg D. *Science.* 2000; 288:1418. [PubMed: 10827951]
32. He Y, Fang J, Taatjes DJ, Nogales E. *Nature.* 2013; 495:481. [PubMed: 23446344]
33. Sikorski TW, Buratowski S. *Curr Opin Cell Biol.* 2009; 21:344. [PubMed: 19411170]
34. Dvir A. *Biochimica et biophysica acta.* 2002; 1577:208. [PubMed: 12213653]
35. (a) Kornberg RD. *Trends Biochem Sci.* 2005; 30:235. [PubMed: 15896740] (b) Conaway RC, Conaway JW. *Semin Cell Dev Biol.* 2011; 22:729. [PubMed: 21821140] (c) Fondell JD. *Biochimica et biophysica acta.* 2013; 1830:3867. [PubMed: 22402254]
36. Bernecky C, Grob P, Ebmeier CC, Nogales E, Taatjes DJ. *PLoS Biol.* 2011; 9:e1000603. [PubMed: 21468301]
37. (a) Ebmeier CC, Taatjes DJ. *Proc Natl Acad Sci U S A.* 2010; 107:11283. [PubMed: 20534441] (b) Kagey MH, Newman JJ, Bilodeau S, Zhan Y, Orlando DA, van Berkum NL, Ebmeier CC, Goossens J, Rahl PB, Levine SS, Taatjes DJ, Dekker J, Young RA. *Nature.* 2010; 467:430. [PubMed: 20720539] (c) Malik S, Roeder RG. *Nature reviews Genetics.* 2010; 11:761.
38. Hu X, Malik S, Negroiu CC, Hubbard K, Velalar CN, Hampton B, Grosu D, Catalano J, Roeder RG, Gnatt A. *Proc Natl Acad Sci U S A.* 2006; 103:9506. [PubMed: 16769904]
39. (a) Cheng B, Li T, Rahl PB, Adamson TE, Loudas NB, Guo J, Varzavand K, Cooper JJ, Hu X, Gnatt A, Young RA, Price DH. *Mol Cell.* 2012; 45:38. [PubMed: 22244331] (b) Jishage M, Malik S, Wagner U, Uberheide B, Ishihama Y, Hu X, Chait BT, Gnatt A, Ren B, Roeder RG. *Mol Cell.* 2012; 45:51. [PubMed: 22244332] (c) Wu YM, Chang JW, Wang CH, Lin YC, Wu PL, Huang SH, Chang CC, Hu X, Gnatt A, Chang WH. *Embo J.* 2012
40. Maldonado E, Drapkin R, Reinberg D. *Methods Enzymol.* 1996; 274:72. [PubMed: 8902797]
41. (a) Cheng B, Price DH. *Methods.* 2009; 48:346. [PubMed: 19409997] (b) Adamson TE, Shore SM, Price DH. *Methods Enzymol.* 2003; 371:264. [PubMed: 14712706]

42. An W, Roeder RG. *Methods Enzymol.* 2004; 377:460. [PubMed: 14979045]
43. Ujvari A, Pal M, Luse DS. *J Biol Chem.* 2011; 286:23160. [PubMed: 21566144]
44. (a) Izban MG, Samkurashvili I, Luse DS. *J Biol Chem.* 1995; 270:2290. [PubMed: 7836462] (b) Chen Y, Yamaguchi Y, Tsugeno Y, Yamamoto J, Yamada T, Nakamura M, Hisatake K, Handa H. *Genes Dev.* 2009; 23:2765. [PubMed: 19952111]
45. (a) Dignam JD, Lebovitz RM, Roeder RG. *Nucleic Acids Res.* 1983; 11:1475. [PubMed: 6828386] (b) Price DH, Sluder AE, Greenleaf AL. *J Biol Chem.* 1987; 262:3244. [PubMed: 3818640]
46. Cheng B, Price DH. *J Biol Chem.* 2007; 282:21901. [PubMed: 17548348]
47. (a) Cheng B, Price DH. *Nucleic Acids Res.* 2008; 36:e135. [PubMed: 18832375] (b) Missra A, Gilmour DS. *Proc Natl Acad Sci U S A.* 2010; 107:11301. [PubMed: 20534440]
48. Furey TS. *Nature reviews Genetics.* 2012; 13:840.
49. Karolchik, D.; Hinrichs, AS.; Kent, WJ. *Current protocols in bioinformatics/editorial board.* Baxevanis, Andreas D., et al., editors. Vol. Chapter 1. 2012.
50. Wang Z, Gerstein M, Snyder M. *Nature reviews Genetics.* 2009; 10:57.
51. (a) Sauliere J, Murigneux V, Wang Z, Marquet E, Barbosa I, Le Tonqueze O, Audic Y, Paillard L, Roest Crollius H, Le Hir H. *Nat Struct Mol Biol.* 2012; 19:1124. [PubMed: 23085716] (b) Licatalosi DD, Mele A, Fak JJ, Ule J, Kayikci M, Chi SW, Clark TA, Schweitzer AC, Blume JE, Wang X, Darnell JC, Darnell RB. *Nature.* 2008; 456:464. [PubMed: 18978773]
52. (a) Hafner M, Landthaler M, Burger L, Khorshid M, Hausser J, Berninger P, Rothballer A, Ascano M Jr, Jungkamp AC, Munschauer M, Ulrich A, Wardle GS, Dewell S, Zavolan M, Tuschl T. *Cell.* 2010; 141:129. [PubMed: 20371350] (b) Hafner M, Landthaler M, Burger L, Khorshid M, Hausser J, Berninger P, Rothballer A, Ascano M, Jungkamp AC, Munschauer M, Ulrich A, Wardle GS, Dewell S, Zavolan M, Tuschl T. *Journal of visualized experiments : JoVE.* 2010
53. Core LJ, Waterfall JJ, Lis JT. *Science.* 2008; 322:1845. [PubMed: 19056941]
54. Core LJ, Waterfall JJ, Gilchrist DA, Fargo DC, Kwak H, Adelman K, Lis JT. *Cell reports.* 2012; 2:1025. [PubMed: 23062713]
55. Kwak H, Fuda NJ, Core LJ, Lis JT. *Science.* 2013; 339:950. [PubMed: 23430654]
56. Churchman LS, Weissman JS. *Nature.* 2011; 469:368. [PubMed: 21248844]
57. Martinez-Rucobo FW, Cramer P. *Biochimica et biophysica acta.* 2013; 1829:9. [PubMed: 22982352]
58. Nudler E, Mustaev A, Lukhtanov E, Goldfarb A. *Cell.* 1997; 89:33. [PubMed: 9094712]
59. Adamson TE, Price DH. *Mol Cell Biol.* 2003; 23:4046. [PubMed: 12773550]
60. Renner DB, Yamaguchi Y, Wada T, Handa H, Price DH. *J Biol Chem.* 2001; 276:42601. [PubMed: 11553615]
61. Price DH, Sluder AE, Greenleaf AL. *Mol Cell Biol.* 1989; 9:1465. [PubMed: 2725511]
62. Rougvie AE, Lis JT. *Mol Cell Biol.* 1990; 10:6041. [PubMed: 2172790]
63. Marshall NF, Price DH. *Mol Cell Biol.* 1992; 12:2078. [PubMed: 1569941]
64. Nechaev S, Adelman K. *Cell Cycle.* 2008; 7:1539. [PubMed: 18469524]
65. (a) Rhee HS, Pugh BF. *Nature.* 2012; 483:295. [PubMed: 22258509] (b) Kim H, Erickson B, Luo W, Seward D, Graber JH, Pollock DD, Megee PC, Bentley DL. *Nat Struct Mol Biol.* 2010; 17:1279. [PubMed: 20835241] (c) Mayer A, Lidschreiber M, Siebert M, Leike K, Soding J, Cramer P. *Nat Struct Mol Biol.* 2010
66. Gaertner B, Johnston J, Chen K, Wallaschek N, Paulson A, Garruss AS, Gaudenz K, De Kumar B, Krumlauf R, Zeitlinger J. *Cell reports.* 2012; 2:1670. [PubMed: 23260668]
67. (a) Seila AC, Calabrese JM, Levine SS, Yeo GW, Rahl PB, Flynn RA, Young RA, Sharp PA. *Science.* 2008; 322:1849. [PubMed: 19056940] (b) He Y, Vogelstein B, Velculescu VE, Papadopoulos N, Kinzler KW. *Science.* 2008; 322:1855. [PubMed: 19056939] (c) Preker P, Nielsen J, Kammiller S, Lykke-Andersen S, Christensen MS, Mapendano CK, Schierup MH, Jensen TH. *Science.* 2008; 322:1851. [PubMed: 19056938]
68. Flynn RA, Almada AE, Zamudio JR, Sharp PA. *Proc Natl Acad Sci U S A.* 2011; 108:10460. [PubMed: 21670248]

69. Rahl PB, Lin CY, Seila AC, Flynn RA, McCuine S, Burge CB, Sharp PA, Young RA. *Cell*. 2010; 141:432. [PubMed: 20434984]
70. Almada AE, Wu X, Kriz AJ, Burge CB, Sharp PA. *Nature*. 2013
71. Sigova AA, Mullen AC, Molinie B, Gupta S, Orlando DA, Guenther MG, Almada AE, Lin C, Sharp PA, Giallourakis CC, Young RA. *Proc Natl Acad Sci U S A*. 2013; 110:2876. [PubMed: 23382218]
72. (a) Xu Z, Wei W, Gagneur J, Perocchi F, Clauder-Munster S, Camblong J, Guffanti E, Stutz F, Huber W, Steinmetz LM. *Nature*. 2009; 457:1033. [PubMed: 19169243] (b) Neil H, Malabat C, d'Aubenton-Carafa Y, Xu Z, Steinmetz LM, Jacquier A. *Nature*. 2009; 457:1038. [PubMed: 19169244]
73. FitzGerald PC, Sturgill D, Shyakhtenko A, Oliver B, Vinson C. *Genome biology*. 2006; 7:R53. [PubMed: 16827941]
74. (a) Yamaguchi Y, Shibata H, Handa H. *Biochimica et biophysica acta*. 2013; 1829:98. [PubMed: 23202475] (b) Hartzog GA, Fu J. *Biochimica et biophysica acta*. 2013; 1829:105. [PubMed: 22982195] (c) Tomar SK, Artsimovitch I. *Chemical reviews*. 2013(d) Werner F. *J Mol Biol*. 2012; 417:13. [PubMed: 22306403]
75. Egyhazi E, Ossoinak A, Pigon A, Holmgren C, Lee JM, Greenleaf AL. *Chromosoma*. 1996; 104:422. [PubMed: 8601337]
76. Sehgal PB, Darnell JE Jr, Tamm I. *Cell*. 1976; 9:473. [PubMed: 1086720]
77. Chodosh LA, Fire A, Samuels M, Sharp PA. *J Biol Chem*. 1989; 264:2250. [PubMed: 2914905]
78. Wada T, Takagi T, Yamaguchi Y, Ferdous A, Imai T, Hirose S, Sugimoto S, Yano K, Hartzog GA, Winston F, Buratowski S, Handa H. *Genes Dev*. 1998; 12:343. [PubMed: 9450929]
79. Yamaguchi Y, Takagi T, Wada T, Yano K, Furuya A, Sugimoto S, Hasegawa J, Handa H. *Cell*. 1999; 97:41. [PubMed: 10199401]
80. Wu CH, Yamaguchi Y, Benjamin LR, Horvat-Gordon M, Washinsky J, Enerly E, Larsson J, Lambertsson A, Handa H, Gilmour D. *Genes Dev*. 2003; 17:1402. [PubMed: 12782658]
81. Fujinaga K, Irwin D, Huang Y, Taube R, Kurosu T, Peterlin BM. *Mol Cell Biol*. 2004; 24:787. [PubMed: 14701750]
82. Yamada T, Yamaguchi Y, Inukai N, Okamoto S, Mura T, Handa H. *Mol Cell*. 2006; 21:227. [PubMed: 16427012]
83. (a) Li J, Gilmour DS. *Embo J*. 2013(b) Li J, Liu Y, Rhee S, Ghosh SKB, Bai L, Pugh F, Gilmour DS. *Mol Cell*. 2013; 50:711. [PubMed: 23746353]
84. Natori S, Takeuchi K, Mizuno D. *Journal of biochemistry*. 1973; 74:1177. [PubMed: 4798604]
85. Reines D, Mote J Jr. *Proc Natl Acad Sci U S A*. 1993; 90:1917. [PubMed: 8446609]
86. Fish RN, Kane CM. *Biochimica et biophysica acta*. 2002; 1577:287. [PubMed: 12213659]
87. Adelman K, Marr MT, Werner J, Saunders A, Ni Z, Andrulis ED, Lis JT. *Mol Cell*. 2005; 17:103. [PubMed: 15629721]
88. Nechaev S, Fargo DC, dos Santos G, Liu L, Gao Y, Adelman K. *Science*. 2010; 327:335. [PubMed: 20007866]
89. Gilmour DS, Lis JT. *Mol Cell Biol*. 1986; 6:3984. [PubMed: 3099167]
90. (a) Lagha M, Bothma JP, Esposito E, Ng S, Stefanik L, Tsui C, Johnston J, Chen K, Gilmour DS, Zeitlinger J, Levine MS. *Cell*. 2013; 153:976. [PubMed: 23706736] (b) Boettiger AN, Levine M. *Science*. 2009; 325:471. [PubMed: 19628867]
91. Gilchrist DA, Nechaev S, Lee C, Ghosh SK, Collins JB, Li L, Gilmour DS, Adelman K. *Genes Dev*. 2008; 22:1921. [PubMed: 18628398]
92. Chopra VS, Cande J, Hong JW, Levine M. *Genes Dev*. 2009; 23:1505. [PubMed: 19515973]
93. Zhu Y, Pe'ery T, Peng J, Ramanathan Y, Marshall N, Marshall T, Amendt B, Mathews MB, Price DH. *Genes Dev*. 1997; 11:2622. [PubMed: 9334325]
94. Peng J, Marshall NF, Price DH. *J Biol Chem*. 1998; 273:13855. [PubMed: 9593731]
95. Peng J, Zhu Y, Milton JT, Price DH. *Genes Dev*. 1998; 12:755. [PubMed: 9499409]
96. (a) Liu Y, Warfield L, Zhang C, Luo J, Allen J, Lang WH, Ranish J, Shokat KM, Hahn S. *Mol Cell Biol*. 2009; 29:4852. [PubMed: 19581288] (b) Keogh MC, Podolny V, Buratowski S. *Mol Cell*

- Biol. 2003; 23:7005. [PubMed: 12972617] (c) Qiu H, Hu C, Gaur NA, Hinnebusch AG. *Embo J*. 2012; 31:3494. [PubMed: 22796944]
97. Peng J, Liu M, Marion J, Zhu Y, Price DH. *Cold Spring Harb Symp Quant Biol*. 1998; 63:365. [PubMed: 10384301]
 98. Sehgal PB, Derman E, Molloy GR, Tamm I, Darnell JE. *Science*. 1976; 194:431. [PubMed: 982026]
 99. Chao SH, Price DH. *J Biol Chem*. 2001; 276:31793. [PubMed: 11431468]
 100. Maiuri P, Knezevich A, De Marco A, Mazza D, Kula A, McNally JG, Marcello A. *EMBO Rep*. 2011
 101. Singh J, Padgett RA. *Nat Struct Mol Biol*. 2009; 16:1128. [PubMed: 19820712]
 102. Wada Y, Ohta Y, Xu M, Tsutsumi S, Minami T, Inoue K, Komura D, Kitakami J, Oshida N, Papantonis A, Izumi A, Kobayashi M, Meguro H, Kanki Y, Mimura I, Yamamoto K, Matakai C, Hamakubo T, Shirahige K, Aburatani H, Kimura H, Kodama T, Cook PR, Ihara S. *Proc Natl Acad Sci U S A*. 2009; 106:18357. [PubMed: 19826084]
 103. (a) Luse DS, Studitsky VM. *RNA biology*. 2011; 8:581. [PubMed: 21519186] (b) Bintu L, Kopaczynska M, Hodges C, Lubkowska L, Kashlev M, Bustamante C. *Nat Struct Mol Biol*. 2011; 18:1394. [PubMed: 22081017] (c) Petesch SJ, Lis JT. *Trends in genetics : TIG*. 2012; 28:285. [PubMed: 22465610] (d) Bondarenko VA, Steele LM, Ujvari A, Gaykalova DA, Kulaeva OI, Polikanov YS, Luse DS, Studitsky VM. *Mol Cell*. 2006; 24:469. [PubMed: 17081995] (e) Yen K, Vinayachandran V, Batta K, Koerber RT, Pugh BF. *Cell*. 2012; 149:1461. [PubMed: 22726434] (f) Kulaeva OI, Hsieh FK, Chang HW, Luse DS, Studitsky VM. *Biochimica et biophysica acta*. 2013; 1829:76. [PubMed: 22982194]
 104. Shilatifard A, Lane WS, Jackson KW, Conaway RC, Conaway JW. *Science*. 1996; 271:1873. [PubMed: 8596958]
 105. Lin C, Smith ER, Takahashi H, Lai KC, Martin-Brown S, Florens L, Washburn MP, Conaway JW, Conaway RC, Shilatifard A. *Mol Cell*. 2010; 37:429. [PubMed: 20159561]
 106. Biswas D, Milne TA, Basrur V, Kim J, Elenitoba-Johnson KS, Allis CD, Roeder RG. *Proc Natl Acad Sci U S A*. 2011; 108:15751. [PubMed: 21896721]
 107. Lin C, Garrett AS, De Kumar B, Smith ER, Gogol M, Seidel C, Krumlauf R, Shilatifard A. *Genes Dev*. 2011; 25:1486. [PubMed: 21764852]
 108. He N, Chan CK, Sobhian B, Chou S, Xue Y, Liu M, Alber T, Benkirane M, Zhou Q. *Proc Natl Acad Sci U S A*. 2011; 108:E636. [PubMed: 21873227]
 109. Liu M, Hsu J, Chan C, Li Z, Zhou Q. *Mol Cell*. 2012; 46:325. [PubMed: 22483617]
 110. Smith ER, Lin C, Garrett AS, Thornton J, Mohaghegh N, Hu D, Jackson J, Saraf A, Swanson SK, Seidel C, Florens L, Washburn MP, Eissenberg JC, Shilatifard A. *Mol Cell*. 2011; 44:954. [PubMed: 22195968]
 111. Luo Z, Lin C, Guest E, Garrett AS, Mohaghegh N, Swanson S, Marshall S, Florens L, Washburn MP, Shilatifard A. *Mol Cell Biol*. 2012; 32:2608. [PubMed: 22547686]
 112. Lin C, Garruss AS, Luo Z, Guo F, Shilatifard A. *Cell*. 2012
 113. (a) Jaehning JA. *Biochimica et biophysica acta*. 2010; 1799:379. [PubMed: 20060942] (b) Zhu B, Mandal SS, Pham AD, Zheng Y, Erdjument-Bromage H, Batra SK, Tempst P, Reinberg D. *Genes Dev*. 2005; 19:1668. [PubMed: 16024656]
 114. (a) Wade PA, Werel W, Fentzke RC, Thompson NE, Leykam JF, Burgess RR, Jaehning JA, Burton ZF. *Protein Expr Purif*. 1996; 8:85. [PubMed: 8812838] (b) Shi X, Finkelstein A, Wolf AJ, Wade PA, Burton ZF, Jaehning JA. *Mol Cell Biol*. 1996; 16:669. [PubMed: 8552095]
 115. Kim J, Guermah M, Roeder RG. *Cell*. 2010; 140:491. [PubMed: 20178742]
 116. (a) Sopta M, Carthew RW, Greenblatt J. *J Biol Chem*. 1985; 260:10353. [PubMed: 3860504] (b) Burton ZF, Killeen M, Sopta M, Ortolan LG, Greenblatt J. *Mol Cell Biol*. 1988; 8:1602. [PubMed: 3380090]
 117. Luse DS. *Transcription*. 2012; 3:156. [PubMed: 22771986]
 118. Flores O, Maldonado E, Reinberg D. *J Biol Chem*. 1989; 264:8913. [PubMed: 2566609]

119. Chen ZA, Jawhari A, Fischer L, Buchen C, Tahir S, Kamenski T, Rasmussen M, Lariviere L, Bukowski-Wills JC, Nilges M, Cramer P, Rappsilber J. *Embo J*. 2010; 29:717. [PubMed: 20094031]
120. Cojocaru M, Jeronimo C, Forget D, Bouchard A, Bergeron D, Cote P, Poirier GG, Greenblatt J, Coulombe B. *The Biochemical journal*. 2008; 409:139. [PubMed: 17848138]
121. Sluder AE, Price DH, Greenleaf AL. *J Biol Chem*. 1988; 263:9917. [PubMed: 2454924]
122. Gu W, Reines D. *J Biol Chem*. 1995; 270:11238. [PubMed: 7744757]
123. Kettenberger H, Armache KJ, Cramer P. *Mol Cell*. 2004; 16:955. [PubMed: 15610738]
124. Sigurdsson S, Dirac-Svejstrup AB, Svejstrup JQ. *Mol Cell*. 2010; 38:202. [PubMed: 20417599]
125. (a) Wang D, Bushnell DA, Huang X, Westover KD, Levitt M, Kornberg RD. *Science*. 2009; 324:1203. [PubMed: 19478184] (b) Cheung AC, Cramer P. *Nature*. 2011; 471:249. [PubMed: 21346759]
126. Palangat M, Renner DB, Price DH, Landick R. *Proc Natl Acad Sci U S A*. 2005; 102:15036. [PubMed: 16214896]
127. Sijbrandi R, Fiedler U, Timmers HT. *Nucleic Acids Res*. 2002; 30:2290. [PubMed: 12034815]
128. Shema E, Kim J, Roeder RG, Oren M. *Mol Cell*. 2011; 42:477. [PubMed: 21596312]
129. Taatjes DJ. *Trends Biochem Sci*. 2010; 35:315. [PubMed: 20299225]
130. Bourbon HM, Aguilera A, Ansari AZ, Asturias FJ, Berk AJ, Bjorklund S, Blackwell TK, Borggreffe T, Carey M, Carlson M, Conaway JW, Conaway RC, Emmons SW, Fondell JD, Freedman LP, Fukasawa T, Gustafsson CM, Han M, He X, Herman PK, Hinnebusch AG, Holmberg S, Holstege FC, Jaehning JA, Kim YJ, Kuras L, Leutz A, Lis JT, Meisterernest M, Naar AM, Nasmyth K, Parvin JD, Ptashne M, Reinberg D, Ronne H, Sadowski I, Sakurai H, Sipiczki M, Sternberg PW, Stillman DJ, Strich R, Struhl K, Svejstrup JQ, Tuck S, Winston F, Roeder RG, Kornberg RD. *Mol Cell*. 2004; 14:553. [PubMed: 15175151]
131. Sato S, Tomomori-Sato C, Parmely TJ, Florens L, Zybailov B, Swanson SK, Banks CA, Jin J, Cai Y, Washburn MP, Conaway JW, Conaway RC. *Mol Cell*. 2004; 14:685. [PubMed: 15175163]
132. Takahashi H, Parmely TJ, Sato S, Tomomori-Sato C, Banks CA, Kong SE, Szutorisz H, Swanson SK, Martin-Brown S, Washburn MP, Florens L, Seidel CW, Lin C, Smith ER, Shilatifard A, Conaway RC, Conaway JW. *Cell*. 2011; 146:92. [PubMed: 21729782]
133. Huang Y, Li W, Yao X, Lin QJ, Yin JW, Liang Y, Heiner M, Tian B, Hui J, Wang G. *Mol Cell*. 2012; 45:459. [PubMed: 22264826]
134. Yin JW, Liang Y, Park JY, Chen D, Yao X, Xiao Q, Liu Z, Jiang B, Fu Y, Bao M, Huang Y, Liu Y, Yan J, Zhu MS, Yang Z, Gao P, Tian B, Li D, Wang G. *Genes Dev*. 2012; 26:2192. [PubMed: 22972934]
135. Hashimoto S, Boissel S, Zarhrate M, Rio M, Munnich A, Egly JM, Colleaux L. *Science*. 2011; 333:1161. [PubMed: 21868677]
136. Zhou H, Spaeth JM, Kim NH, Xu X, Friez MJ, Schwartz CE, Boyer TG. *Proc Natl Acad Sci U S A*. 2012; 109:19763. [PubMed: 23091001]
137. Makinen N, Mehine M, Tolvanen J, Kaasinen E, Li Y, Lehtonen HJ, Gentile M, Yan J, Enge M, Taipale M, Aavikko M, Katainen R, Virolainen E, Bohling T, Koski TA, Launonen V, Sjoberg J, Taipale J, Vahteristo P, Aaltonen LA. *Science*. 2011; 334:252. [PubMed: 21868628]
138. (a) Huang S, Holzel M, Knijnenburg T, Schlicker A, Roepman P, McDermott U, Garnett M, Grenrum W, Sun C, Prahallad A, Groenendijk FH, Mittempergher L, Nijkamp W, Neefjes J, Salazar R, Ten Dijke P, Uramoto H, Tanaka F, Beijersbergen RL, Wessels LF, Bernards R. *Cell*. 2012; 151:937. [PubMed: 23178117] (b) Feliciano P. *Nat Genet*. 2012; 45:11.
139. Knuesel MT, Meyer KD, Donner AJ, Espinosa JM, Taatjes DJ. *Mol Cell Biol*. 2009; 29:650. [PubMed: 19047373]
140. Galbraith MD, Donner AJ, Espinosa JM. *Transcription*. 2010; 1:4. [PubMed: 21327159]
141. Knuesel MT, Meyer KD, Bernecky C, Taatjes DJ. *Genes Dev*. 2009; 23:439. [PubMed: 19240132]
142. Donner AJ, Szostek S, Hoover JM, Espinosa JM. *Mol Cell*. 2007; 27:121. [PubMed: 17612495]

143. Donner AJ, Ebmeier CC, Taatjes DJ, Espinosa JM. *Nat Struct Mol Biol.* 2010; 17:194. [PubMed: 20098423]
144. (a) Li T, Price D. *Transcription.* 2012; 3:177. [PubMed: 22771989] (b) Espinosa JM. *Mol Cell.* 2012; 45:3. [PubMed: 22244325]
145. Gilchrist DA, Dos Santos G, Fargo DC, Xie B, Gao Y, Li L, Adelman K. *Cell.* 2010; 143:540. [PubMed: 21074046]
146. Gilchrist DA, Adelman K. *Biochimica et biophysica acta.* 2012; 1819:700. [PubMed: 22406341]
147. Guermah M, Kim J, Roeder RG. *Methods.* 2009; 48:353. [PubMed: 19272450]
148. Luse DS, Spangler LC, Ujvari A. *J Biol Chem.* 2011; 286:6040. [PubMed: 21177855]
149. (a) Orphanides G, Wu WH, Lane WS, Hampsey M, Reinberg D. *Nature.* 1999; 400:284. [PubMed: 10421373] (b) Belotserkovskaya R, Oh S, Bondarenko VA, Orphanides G, Studitsky VM, Reinberg D. *Science.* 2003; 301:1090. [PubMed: 12934006] (c) Hsieh FK, Kulaeva OI, Patel SS, Dyer PN, Luger K, Reinberg D, Studitsky VM. *Proc Natl Acad Sci U S A.* 2013(d) Orphanides G, LeRoy G, Chang CH, Luse DS, Reinberg D. *Cell.* 1998; 92:105. [PubMed: 9489704]
150. (a) Smolle M, Workman JL. *Biochimica et biophysica acta.* 2013; 1829:84. [PubMed: 22982198] (b) Selth LA, Sigurdsson S, Svejstrup JQ. *Annu Rev Biochem.* 2010; 79:271. [PubMed: 20367031]
151. Petesch SJ, Lis JT. *Cell.* 2008; 134:74. [PubMed: 18614012]
152. Orphanides G, Reinberg D. *Nature.* 2000; 407:471. [PubMed: 11028991]
153. Schwabish MA, Struhl K. *Mol Cell Biol.* 2004; 24:10111. [PubMed: 15542822]
154. Kristjuhan A, Svejstrup JQ. *Embo J.* 2004; 23:4243. [PubMed: 15457216]
155. Teves SS, Henikoff S. *Biochemistry and cell biology = Biochimie et biologie cellulaire.* 2013; 91:42. [PubMed: 23442140]
156. (a) Li Q, Price JP, Byers SA, Cheng D, Peng J, Price DH. *J Biol Chem.* 2005; 280:28819. [PubMed: 15965233] (b) Chen R, Liu M, Li H, Xue Y, Ramey WN, He N, Ai N, Luo H, Zhu Y, Zhou N, Zhou Q. *Genes Dev.* 2008; 22:1356. [PubMed: 18483222] (c) Chen R, Yang Z, Zhou Q. *J Biol Chem.* 2004; 279:4153. [PubMed: 14627702]
157. (a) Nguyen VT, Kiss T, Michels AA, Bensaude O. *Nature.* 2001; 414:322. [PubMed: 11713533] (b) Yang Z, Zhu Q, Luo K, Zhou Q. *Nature.* 2001; 414:317. [PubMed: 11713532]
158. Peterlin BM, Brogie JE, Price DH. *Wiley interdisciplinary reviews. RNA.* 2012; 3:92. [PubMed: 21853533]
159. Nguyen D, Krueger BJ, Sedore SC, Brogie JE, Rogers JT, Rajendra TK, Saunders A, Matera AG, Lis JT, Uguen P, Price DH. *Nucleic Acids Res.* 2012; 40:5283. [PubMed: 22379134]
160. (a) Garriga J, Peng J, Parreno M, Price DH, Henderson EE, Grana X. *Oncogene.* 1998; 17:3093. [PubMed: 9872325] (b) Haaland RE, Herrmann CH, Rice AP. *AIDS.* 2003; 17:2429. [PubMed: 14600513]
161. Marz M, Donath A, Verstraete N, Nguyen VT, Stadler PF, Bensaude O. *Mol Biol Evol.* 2009; 26:2821. [PubMed: 19734296]
162. Wassarman DA, Steitz JA. *Mol Cell Biol.* 1991; 11:3432. [PubMed: 1646389]
163. Krueger BJ, Varzavand K, Cooper JJ, Price DH. *PLoS One.* 2010; 5:e12335. [PubMed: 20808803]
164. Diribarne G, Bensaude O. *RNA biology.* 2009; 6:122. [PubMed: 19246988]
165. (a) Byers SA, Price JP, Cooper JJ, Li Q, Price DH. *J Biol Chem.* 2005; 280:16360. [PubMed: 15713662] (b) Yik JH, Chen R, Pezda AC, Zhou Q. *J Biol Chem.* 2005; 280:16368. [PubMed: 15713661] (c) Michels AA, Fraldi A, Li Q, Adamson TE, Bonnet F, Nguyen VT, Sedore SC, Price JP, Price DH, Lania L, Bensaude O. *Embo J.* 2004; 23:2608. [PubMed: 15201869]
166. Krueger BJ, Jeronimo C, Roy BB, Bouchard A, Barrandon C, Byers SA, Searcey CE, Cooper JJ, Bensaude O, Cohen EA, Coulombe B, Price DH. *Nucleic Acids Res.* 2008; 36:2219. [PubMed: 18281698]
167. Peterlin BM, Price DH. *Mol Cell.* 2006; 23:297. [PubMed: 16885020]
168. (a) Blazek D, Barboric M, Kohoutek J, Oven I, Peterlin BM. *Nucleic Acids Res.* 2005; 33:7000. [PubMed: 16377779] (b) Dulac C, Michels AA, Fraldi A, Bonnet F, Nguyen VT, Napolitano G,

- Lania L, Bensaude O. *J Biol Chem*. 2005; 280:30619. [PubMed: 15994294] (c) Schulte A, Czudnochowski N, Barboric M, Schonichen A, Blazek D, Peterlin BM, Geyer M. *J Biol Chem*. 2005; 280:24968. [PubMed: 15855166]
169. (a) Yik JH, Chen R, Nishimura R, Jennings JL, Link AJ, Zhou Q. *Mol Cell*. 2003; 12:971. [PubMed: 14580347] (b) Michels AA, Nguyen VT, Fraldi A, Labas V, Edwards M, Bonnet F, Lania L, Bensaude O. *Mol Cell Biol*. 2003; 23:4859. [PubMed: 12832472]
170. Li Q, Cooper JJ, Altwerger GH, Feldkamp MD, Shea MA, Price DH. *Nucleic Acids Res*. 2007; 35:2503. [PubMed: 17395637]
171. Sedore SC, Byers SA, Biglione S, Price JP, Maury WJ, Price DH. *Nucleic Acids Res*. 2007; 35:4347. [PubMed: 17576689]
172. Fujimoto Y, Nakamura Y, Ohuchi S. *Biochimie*. 2012; 94:1900. [PubMed: 22609015]
173. Lebars I, Martinez-Zapien D, Durand A, Coutant J, Kieffer B, Dock-Bregeon AC. *Nucleic Acids Res*. 2010; 38:7749. [PubMed: 20675720]
174. Egloff S, Van Herreweghe E, Kiss T. *Mol Cell Biol*. 2006; 26:630. [PubMed: 16382153]
175. Belanger F, Baigude H, Rana TM. *J Mol Biol*. 2009; 386:1094. [PubMed: 19244621]
176. Fujinaga K, Barboric M, Li Q, Luo Z, Price DH, Peterlin BM. *Nucleic Acids Res*. 2012; 40:9160. [PubMed: 22821562]
177. (a) Yoshikawa N, Shimizu N, Sano M, Ohnuma K, Iwata S, Hosono O, Fukuda K, Morimoto C, Tanaka H. *Biochem Biophys Res Commun*. 2008; 371:44. [PubMed: 18407829] (b) Shimizu N, Ouchida R, Yoshikawa N, Hisada T, Watanabe H, Okamoto K, Kusuhara M, Handa H, Morimoto C, Tanaka H. *Proc Natl Acad Sci U S A*. 2005; 102:8555. [PubMed: 15941832]
178. Mascareno EJ, Belashov I, Siddiqui MA, Liu F, Dhar-Mascareno M. *The Prostate*. 2012; 72:1035. [PubMed: 22095517]
179. (a) Lew QJ, Chia YL, Chu KL, Lam YT, Gurumurthy M, Xu S, Lam KP, Cheong N, Chao SH. *J Biol Chem*. 2012; 287:36443. [PubMed: 22948151] (b) Lau J, Lew QJ, Diribarne G, Michels AA, Dey A, Bensaude O, Lane DP, Chao SH. *Cell Cycle*. 2009; 8:2247. [PubMed: 19617712]
180. Kusuhara MNK, Kimura K, Maass N, Manabe T, Ishikawa S, Aikawa M, Miyazaki K, Yamaguchi K. *Biomed Res*. 1999; 20:273.
181. Ghatpande S, Goswami S, Mathew S, Rong G, Cai L, Shafiq S, Siddiqui MA. *Dev Biol*. 1999; 208:210. [PubMed: 10075853]
182. (a) Espinoza-Derout J, Wagner M, Shahmiri K, Mascareno E, Chaqour B, Siddiqui MA. *Cardiovasc Res*. 2007; 75:129. [PubMed: 17459355] (b) Dey A, Chao SH, Lane DP. *Cell Cycle*. 2007; 6:1856. [PubMed: 17671421] (c) Sano M, Schneider MD. *Circ Res*. 2004; 95:867. [PubMed: 15514168]
183. Yoshikawa N, Shimizu N, Maruyama T, Sano M, Matsuhashi T, Fukuda K, Kataoka M, Satoh T, Ojima H, Sawai T, Morimoto C, Kuribara A, Hosono O, Tanaka H. *PLoS One*. 2012; 7:e52522. [PubMed: 23300697]
184. (a) Galatioto J, Mascareno E, Siddiqui MA. *J Cell Sci*. 2010; 123:3789. [PubMed: 20940258] (b) Hong P, Chen K, Huang B, Liu M, Cui M, Rozenberg I, Chaqour B, Pan X, Barton ER, Jiang XC, Siddiqui MA. *The Journal of clinical investigation*. 2012; 122:3873. [PubMed: 23023707] (c) Mascareno E, Galatioto J, Rozenberg I, Saliccioli L, Kamran H, Lazar JM, Liu F, Pedrazzini T, Siddiqui MA. *J Biol Chem*. 2012; 287:13084. [PubMed: 22308025]
185. (a) Ketchart W, Smith KM, Krupka T, Wittmann BM, Hu Y, Rayman PA, Doughman YQ, Albert JM, Bai X, Finke JH, Xu Y, Exner AA, Montano MM. *Oncogene*. 2012(b) Ketchart W, Ogba N, Kresak A, Albert JM, Pink JJ, Montano MM. *Oncogene*. 2011; 30:3563. [PubMed: 21423213]
186. Gupta S, Busch RK, Singh R, Reddy R. *J Biol Chem*. 1990; 265:19137. [PubMed: 2229067]
187. Jeronimo C, Forget D, Bouchard A, Li Q, Chua G, Poitras C, Therien C, Bergeron D, Bourassa S, Greenblatt J, Chabot B, Poirier GG, Hughes TR, Blanchette M, Price DH, Coulombe B. *Mol Cell*. 2007; 27:262. [PubMed: 17643375]
188. Xue Y, Yang Z, Chen R, Zhou Q. *Nucleic Acids Res*. 2010; 38:360. [PubMed: 19906723]
189. Cosgrove MS, Ding Y, Rennie WA, Lane MJ, Hanes SD. *Wiley interdisciplinary reviews. RNA*. 2012; 3:633. [PubMed: 22740346]

190. (a) He N, Jahchan NS, Hong E, Li Q, Bayfield MA, Maraia RJ, Luo K, Zhou Q. *Mol Cell*. 2008; 29:588. [PubMed: 18249148] (b) Markert A, Grimm M, Martinez J, Wiesner J, Meyerhans A, Meyuhas O, Sickmann A, Fischer U. *EMBO Rep*. 2008; 9:569. [PubMed: 18483487]
191. Barboric M, Lenasi T, Chen H, Johansen EB, Guo S, Peterlin BM. *Proc Natl Acad Sci U S A*. 2009; 106:7798. [PubMed: 19416841]
192. Bayfield MA, Yang R, Maraia RJ. *Biochimica et biophysica acta*. 2010; 1799:365. [PubMed: 20138158]
193. (a) Alazami AM, Al-Owain M, Alzahrani F, Shuaib T, Al-Shamrani H, Al-Falki YH, Al-Qahtani SM, Alsheddi T, Colak D, Alkuraya FS. *Hum Mutat*. 2012; 33:1429. [PubMed: 22865833] (b) Cheng Y, Jin Z, Agarwal R, Ma K, Yang J, Ibrahim S, Olaru AV, David S, Ashktorab H, Smoot DT, Duncan MD, Hutcheon DF, Abraham JM, Meltzer SJ, Mori Y. *Laboratory investigation; a journal of technical methods and pathology*. 2012; 92:1013.
194. (a) Barrandon C, Bonnet F, Nguyen VT, Labas V, Bensaude O. *Mol Cell Biol*. 2007; 27:6996. [PubMed: 17709395] (b) Van Herreweghe E, Egloff S, Goiffon I, Jady BE, Froment C, Monsarrat B, Kiss T. *Embo J*. 2007; 26:3570. [PubMed: 17611602]
195. Tahirov TH, Babayeva ND, Varzavand K, Cooper JJ, Sedore SC, Price DH. *Nature*. 2010; 465:747. [PubMed: 20535204]
196. Yang Z, Yik JH, Chen R, He N, Jang MK, Ozato K, Zhou Q. *Mol Cell*. 2005; 19:535. [PubMed: 16109377]
197. Casse C, Giannoni F, Nguyen VT, Dubois MF, Bensaude O. *J Biol Chem*. 1999; 274:16097. [PubMed: 10347161]
198. Bartholomeeusen K, Xiang Y, Fujinaga K, Peterlin BM. *J Biol Chem*. 2012; 287:36609. [PubMed: 22952229]
199. Kim YJ, Greer CB, Cecchini KR, Harris LN, Tuck DP, Kim TH. *Oncogene*. 2013; 32:2828. [PubMed: 23435418]
200. Biglione S, Byers SA, Price JP, Nguyen VT, Bensaude O, Price DH, Maury W. *Retrovirology*. 2007; 4:47. [PubMed: 17625008]
201. Jang MK, Mochizuki K, Zhou M, Jeong HS, Brady JN, Ozato K. *Mol Cell*. 2005; 19:523. [PubMed: 16109376]
202. D'Orso I, Frankel AD. *Nat Struct Mol Biol*. 2010; 17:815. [PubMed: 20562857]
203. Prasanth KV, Camiolo M, Chan G, Tripathi V, Denis L, Nakamura T, Hubner MR, Spector DL. *Molecular biology of the cell*. 2010; 21:4184. [PubMed: 20881057]
204. Meinhart A, Kamenski T, Hoepfner S, Baumli S, Cramer P. *Genes Dev*. 2005; 19:1401. [PubMed: 15964991]
205. Buratowski S. *Mol Cell*. 2009; 36:541. [PubMed: 19941815]
206. Heidemann M, Hintermair C, Voss K, Eick D. *Biochimica et biophysica acta*. 2013; 1829:55. [PubMed: 22982363]
207. (a) Baskaran R, Dahmus ME, Wang JY. *Proc Natl Acad Sci U S A*. 1993; 90:11167. [PubMed: 7504297] (b) Mayer A, Heidemann M, Lidschreiber M, Schreieck A, Sun M, Hintermair C, Kremmer E, Eick D, Cramer P. *Science*. 2012; 336:1723. [PubMed: 22745433]
208. Baskaran R, Chiang GG, Mysliwiec T, Kruh GD, Wang JY. *J Biol Chem*. 1997; 272:18905. [PubMed: 9228069]
209. (a) Blazek D, Kohoutek J, Bartholomeeusen K, Johansen E, Hulinkova P, Luo Z, Cimermanic P, Ule J, Peterlin BM. *Genes Dev*. 2011; 25:2158. [PubMed: 22012619] (b) Bartkowiak B, Liu P, Phatnani HP, Fuda NJ, Cooper JJ, Price DH, Adelman K, Lis JT, Greenleaf AL. *Genes Dev*. 2010; 24:2303. [PubMed: 20952539]
210. Hintermair C, Heidemann M, Koch F, Descostes N, Gut M, Gut I, Fenouil R, Ferrier P, Flatley A, Kremmer E, Chapman RD, Andrau JC, Eick D. *Embo J*. 2012; 31:2784. [PubMed: 22549466]
211. (a) Akhtar MS, Heidemann M, Tietjen JR, Zhang DW, Chapman RD, Eick D, Ansari AZ. *Mol Cell*. 2009; 34:387. [PubMed: 19450536] (b) Glover-Cutter K, Larochelle S, Erickson B, Zhang C, Shokat K, Fisher RP, Bentley DL. *Mol Cell Biol*. 2009; 29:5455. [PubMed: 19667075] (c) Kim M, Suh H, Cho EJ, Buratowski S. *J Biol Chem*. 2009; 284:26421. [PubMed: 19679665]

212. (a) Liao SM, Zhang J, Jeffery DA, Koleske AJ, Thompson CM, Chao DM, Viljoen M, van Vuuren HJ, Young RA. *Nature*. 1995; 374:193. [PubMed: 7877695] (b) Sun X, Zhang Y, Cho H, Rickert P, Lees E, Lane W, Reinberg D. *Mol Cell*. 1998; 2:213. [PubMed: 9734358]
213. Dahmus ME. *J Biol Chem*. 1996; 271:19009. [PubMed: 8759772]
214. Chesnut JD, Stephens JH, Dahmus ME. *J Biol Chem*. 1992; 267:10500. [PubMed: 1316903]
215. Cho H, Kim TK, Mancebo H, Lane WS, Flores O, Reinberg D. *Genes Dev*. 1999; 13:1540. [PubMed: 10385623]
216. Lin PS, Dubois MF, Dahmus ME. *J Biol Chem*. 2002; 277:45949. [PubMed: 12351650]
217. (a) Ghosh A, Shuman S, Lima CD. *Mol Cell*. 2008; 32:478. [PubMed: 19026779] (b) Hausmann S, Shuman S. *J Biol Chem*. 2002; 277:21213. [PubMed: 11934898] (c) Cho EJ, Kobor MS, Kim M, Greenblatt J, Buratowski S. *Genes Dev*. 2001; 15:3319. [PubMed: 11751637]
218. Sun ZW, Hampsey M. *Mol Cell Biol*. 1996; 16:1557. [PubMed: 8657130]
219. (a) Krishnamurthy S, He X, Reyes-Reyes M, Moore C, Hampsey M. *Mol Cell*. 2004; 14:387. [PubMed: 15125841] (b) Xiang K, Nagaike T, Xiang S, Kilic T, Beh MM, Manley JL, Tong L. *Nature*. 2010
220. (a) Zhang DW, Mosley AL, Ramisetty SR, Rodriguez-Molina JB, Washburn MP, Ansari AZ. *J Biol Chem*. 2012; 287:8541. [PubMed: 22235117] (b) Xiang K, Manley JL, Tong L. *Genes Dev*. 2012; 26:2265. [PubMed: 23070812]
221. (a) Yeo M, Lin PS, Dahmus ME, Gill GN. *J Biol Chem*. 2003; 278:26078. [PubMed: 12721286] (b) Kamenski T, Heilmeyer S, Meinhardt A, Cramer P. *Mol Cell*. 2004; 15:399. [PubMed: 15304220]
222. Zhang Y, Kim Y, Genoud N, Gao J, Kelly JW, Pfaff SL, Gill GN, Dixon JE, Noel JP. *Mol Cell*. 2006; 24:759. [PubMed: 17157258]
223. (a) Mosley AL, Pattenden SG, Carey M, Venkatesh S, Gilmore JM, Florens L, Workman JL, Washburn MP. *Mol Cell*. 2009; 34:168. [PubMed: 19394294] (b) Egloff S, Zaborowska J, Laitem C, Kiss T, Murphy S. *Mol Cell*. 2012; 45:111. [PubMed: 22137580]
224. Xiang K, Manley JL, Tong L. *Nature communications*. 2012; 3:946.
225. Shatkin AJ. *Cell*. 1976; 9:645. [PubMed: 1017010]
226. (a) Rasmussen EB, Lis JT. *Proc Natl Acad Sci U S A*. 1993; 90:7923. [PubMed: 8367444] (b) Moteki S, Price D. *Mol Cell*. 2002; 10:599. [PubMed: 12408827]
227. Shatkin AJ, Manley JL. *Nat Struct Biol*. 2000; 7:838. [PubMed: 11017188]
228. Topisirovic I, Svitkin YV, Sonenberg N, Shatkin AJ. *Wiley interdisciplinary reviews. RNA*. 2011; 2:277. [PubMed: 21957010]
229. Ghosh A, Lima CD. *Wiley interdisciplinary reviews. RNA*. 2010; 1:152. [PubMed: 21956912]
230. Ghosh A, Shuman S, Lima CD. *Mol Cell*. 2011; 43:299. [PubMed: 21683636]
231. (a) McCracken S, Fong N, Rosonina E, Yankulov K, Brothers G, Siderovski D, Hessel A, Foster S, Shuman S, Bentley DL. *Genes Dev*. 1997; 11:3306. [PubMed: 9407024] (b) Cho EJ, Takagi T, Moore CR, Buratowski S. *Genes Dev*. 1997; 11:3319. [PubMed: 9407025]
232. Ho CK, Shuman S. *Mol Cell*. 1999; 3:405. [PubMed: 10198643]
233. Suh MH, Meyer PA, Gu M, Ye P, Zhang M, Kaplan CD, Lima CD, Fu J. *J Biol Chem*. 2010; 285:34027. [PubMed: 20720002]
234. Wen Y, Shatkin AJ. *Genes Dev*. 1999; 13:1774. [PubMed: 10421630]
235. Mandal SS, Chu C, Wada T, Handa H, Shatkin AJ, Reinberg D. *Proc Natl Acad Sci U S A*. 2004; 101:7572. [PubMed: 15136722]
236. (a) Cowling VH, Cole MD. *Genes & cancer*. 2010; 1:576. [PubMed: 21170289] (b) Cole MD, Cowling VH. *Oncogene*. 2009; 28:1169. [PubMed: 19137018]
237. (a) Jiao X, Chang JH, Kilic T, Tong L, Kiledjian M. *Mol Cell*. 2013; 50:104. [PubMed: 23523372] (b) Chang JH, Jiao X, Chiba K, Oh C, Martin CE, Kiledjian M, Tong L. *Nat Struct Mol Biol*. 2012; 19:1011. [PubMed: 22961381] (c) Jiao X, Xiang S, Oh C, Martin CE, Tong L, Kiledjian M. *Nature*. 2010; 467:608. [PubMed: 20802481]
238. (a) Jurica MS, Moore MJ. *Mol Cell*. 2003; 12:5. [PubMed: 12887888] (b) Zhou Z, Licklider LJ, Gygi SP, Reed R. *Nature*. 2002; 419:182. [PubMed: 12226669] (c) Jurica MS, Licklider LJ, Gygi SR, Grigorieff N, Moore MJ. *RNA*. 2002; 8:426. [PubMed: 11991638]

239. (a) Mortillaro MJ, Blencowe BJ, Wei X, Nakayasu H, Du L, Warren SL, Sharp PA, Berezney R. *Proc Natl Acad Sci U S A*. 1996; 93:8253. [PubMed: 8710856] (b) Hirose Y, Tacke R, Manley JL. *Genes Dev*. 1999; 13:1234. [PubMed: 10346811]
240. Misteli T, Spector DL. *Mol Cell*. 1999; 3:697. [PubMed: 10394358]
241. (a) David CJ, Manley JL. *Transcription*. 2011; 2:221. [PubMed: 22231118] (b) David CJ, Boyne AR, Millhouse SR, Manley JL. *Genes Dev*. 2011; 25:972. [PubMed: 21536736]
242. (a) Fong YW, Zhou Q. *Nature*. 2001; 414:929. [PubMed: 11780068] (b) Lin S, Coutinho-Mansfield G, Wang D, Pandit S, Fu XD. *Nat Struct Mol Biol*. 2008; 15:819. [PubMed: 18641664]
243. (a) Montes M, Cloutier A, Sanchez-Hernandez N, Michelle L, Lemieux B, Blanchette M, Hernandez-Munain C, Chabot B, Sune C. *Mol Cell Biol*. 2012; 32:751. [PubMed: 22158966] (b) Ip JY, Schmidt D, Pan Q, Ramani AK, Fraser AG, Odom DT, Blencowe BJ. *Genome Res*. 2011; 21:390. [PubMed: 21163941] (c) Kornblihtt AR. *Adv Exp Med Biol*. 2007; 623:175. [PubMed: 18380347] (d) Nogues G, Kadener S, Cramer P, de la Mata M, Fededa JP, Blaustein M, Srebrow A, Kornblihtt AR. *IUBMB life*. 2003; 55:235. [PubMed: 12880204] (e) Shukla S, Oberdoerffer S. *Biochimica et biophysica acta*. 2012; 1819:673. [PubMed: 22326677]
244. Close P, East P, Dirac-Svejstrup AB, Hartmann H, Heron M, Maslen S, Chariot A, Soding J, Skehel M, Svejstrup JQ. *Nature*. 2012; 484:386. [PubMed: 22446626]
245. (a) Proudfoot N. *Cell*. 1996; 87:779. [PubMed: 8945502] (b) Colgan DF, Manley JL. *Genes Dev*. 1997; 11:2755. [PubMed: 9353246]
246. Shi Y, Di Giammartino DC, Taylor D, Sarkeshik A, Rice WJ, Yates JR 3rd, Frank J, Manley JL. *Mol Cell*. 2009; 33:365. [PubMed: 19217410]
247. (a) Dominski Z, Marzluff WF. *Gene*. 2007; 396:373. [PubMed: 17531405] (b) Dominski Z. *Crit Rev Biochem Mol Biol*. 2007; 42:67. [PubMed: 17453916] (c) Marzluff WF, Wagner EJ, Duronio RJ. *Nature reviews Genetics*. 2008; 9:843. (d) Tan D, Marzluff WF, Dominski Z, Tong L. *Science*. 2013; 339:318. [PubMed: 23329046]
248. (a) Sullivan KD, Steiniger M, Marzluff WF. *Mol Cell*. 2009; 34:322. [PubMed: 19450530] (b) Dominski Z, Yang XC, Marzluff WF. *Cell*. 2005; 123:37. [PubMed: 16213211]
249. Adamson TE, Shutt DC, Price DH. *J Biol Chem*. 2005; 280:32262. [PubMed: 16041059]
250. McCracken S, Fong N, Yankulov K, Ballantyne S, Pan G, Greenblatt J, Patterson SD, Wickens M, Bentley DL. *Nature*. 1997; 385:357. [PubMed: 9002523]
251. (a) Meinhart A, Cramer P. *Nature*. 2004; 430:223. [PubMed: 15241417] (b) Licatalosi DD, Geiger G, Minet M, Schroeder S, Cilli K, McNeil JB, Bentley DL. *Mol Cell*. 2002; 9:1101. [PubMed: 12049745]
252. (a) Ahn SH, Kim M, Buratowski S. *Mol Cell*. 2004; 13:67. [PubMed: 14731395] (b) Ni Z, Schwartz BE, Werner J, Suarez JR, Lis JT. *Mol Cell*. 2004; 13:55. [PubMed: 14731394] (c) Guo J, Garrett M, Micklem G, Brogna S. *Mol Cell Biol*. 2011; 31:639. [PubMed: 21135120] (d) Kim M, Ahn SH, Krogan NJ, Greenblatt JF, Buratowski S. *Embo J*. 2004; 23:354. [PubMed: 14739930]
253. Glover-Cutter K, Kim S, Espinosa J, Bentley DL. *Nat Struct Mol Biol*. 2008; 15:71. [PubMed: 18157150]
254. Dantonel JC, Murthy KG, Manley JL, Tora L. *Nature*. 1997; 389:399. [PubMed: 9311784]
255. El Kaderi B, Medler S, Raghunayakula S, Ansari A. *J Biol Chem*. 2009; 284:25015. [PubMed: 19602510]
256. (a) Sheldon KE, Mauger DM, Arndt KM. *Mol Cell*. 2005; 20:225. [PubMed: 16246725] (b) Rosonina E, Manley JL. *Mol Cell*. 2005; 20:167. [PubMed: 16246718] (c) Penheiter KL, Washburn TM, Porter SE, Hoffman MG, Jaehning JA. *Mol Cell*. 2005; 20:213. [PubMed: 16246724]
257. Nagaike T, Logan C, Hotta I, Rozenblatt-Rosen O, Meyerson M, Manley JL. *Mol Cell*. 2011; 41:409. [PubMed: 21329879]
258. Martincic K, Alkan SA, Cheadle A, Borghesi L, Milcarek C. *Nat Immunol*. 2009; 10:1102. [PubMed: 19749764]
259. Hara R, Selby CP, Liu M, Price DH, Sancar A. *J Biol Chem*. 1999; 274:24779. [PubMed: 10455150]

260. (a) Szalontai T, Gaspar I, Beleczi I, Kerekes I, Erdelyi M, Boros I, Szabad J. *Genetics*. 2009; 181:367. [PubMed: 19047413] (b) Shermoen AW, O'Farrell PH. *Cell*. 1991; 67:303. [PubMed: 1680567]
261. Liu M, Xie Z, Price DH. *J Biol Chem*. 1998; 273:25541. [PubMed: 9748214]
262. (a) Xie Z, Price DH. *J Biol Chem*. 1998; 273:3771. [PubMed: 9452510] (b) Xie Z, Price DH. *J Biol Chem*. 1996; 271:11043. [PubMed: 8626643]
263. Ganesan A, Spivak G, Hanawalt PC. *Progress in molecular biology and translational science*. 2012; 110:25. [PubMed: 22749141]
264. Brannan K, Kim H, Erickson B, Glover-Cutter K, Kim S, Fong N, Kiemele L, Hansen K, Davis R, Lykke-Andersen J, Bentley DL. *Mol Cell*. 2012; 46:311. [PubMed: 22483619]
265. (a) LaCava J, Houseley J, Saveanu C, Petfalski E, Thompson E, Jacquier A, Tollervey D. *Cell*. 2005; 121:713. [PubMed: 15935758] (b) Wyers F, Rougemaille M, Badis G, Rousselle JC, Dufour ME, Boulay J, Regnault B, Devaux F, Namane A, Seraphin B, Libri D, Jacquier A. *Cell*. 2005; 121:725. [PubMed: 15935759]
266. (a) Vasiljeva L, Kim M, Terzi N, Soares LM, Buratowski S. *Mol Cell*. 2008; 29:313. [PubMed: 18280237] (b) Thiebaut M, Kisseleva-Romanova E, Rougemaille M, Boulay J, Libri D. *Mol Cell*. 2006; 23:853. [PubMed: 16973437]
267. (a) Steinmetz EJ, Conrad NK, Brow DA, Corden JL. *Nature*. 2001; 413:327. [PubMed: 11565036] (b) Vasiljeva L, Kim M, Mutschler H, Buratowski S, Meinhart A. *Nat Struct Mol Biol*. 2008; 15:795. [PubMed: 18660819] (c) Gudipati RK, Villa T, Boulay J, Libri D. *Nat Struct Mol Biol*. 2008; 15:786. [PubMed: 18660821] (d) Porrua O, Hobor F, Boulay J, Kubicek K, D'Aubenton-Carafa Y, Gudipati RK, Stefl R, Libri D. *Embo J*. 2012; 31:3935. [PubMed: 23032188]
268. Porrua O, Libri D. *Nat Struct Mol Biol*. 2013
269. Skourti-Stathaki K, Proudfoot NJ, Gromak N. *Mol Cell*. 2011; 42:794. [PubMed: 21700224]
270. (a) Connelly S, Manley JL. *Genes Dev*. 1988; 2:440. [PubMed: 2836265] (b) Proudfoot NJ. *Trends Biochem Sci*. 1989; 14:105. [PubMed: 2658217]
271. (a) West S, Gromak N, Proudfoot NJ. *Nature*. 2004; 432:522. [PubMed: 15565158] (b) Kim M, Krogan NJ, Vasiljeva L, Rando OJ, Nedeia E, Greenblatt JF, Buratowski S. *Nature*. 2004; 432:517. [PubMed: 15565157]
272. (a) Zhang Z, Fu J, Gilmour DS. *Genes Dev*. 2005; 19:1572. [PubMed: 15998810] (b) Zhang Z, Gilmour DS. *Mol Cell*. 2006; 21:65. [PubMed: 16387654]
273. Luo W, Johnson AW, Bentley DL. *Genes Dev*. 2006; 20:954. [PubMed: 16598041]
274. (a) Tan-Wong SM, Zaugg JB, Camblong J, Xu Z, Zhang DW, Mischo HE, Ansari AZ, Luscombe NM, Steinmetz LM, Proudfoot NJ. *Science*. 2012; 338:671. [PubMed: 23019609] (b) Singh BN, Hampsey M. *Mol Cell*. 2007; 27:806. [PubMed: 17803944] (c) O'Sullivan JM, Tan-Wong SM, Morillon A, Lee B, Coles J, Mellor J, Proudfoot NJ. *Nat Genet*. 2004; 36:1014. [PubMed: 15314641]
275. (a) Edelman LB, Fraser P. *Curr Opin Genet Dev*. 2012; 22:110. [PubMed: 22365496] (b) Rieder D, Trajanoski Z, McNally JG. *Frontiers in genetics*. 2012; 3:221. [PubMed: 23109938] (c) Deng B, Melnik S, Cook PR. *Semin Cancer Biol*. 2012
276. (a) Jackson DA, Hassan AB, Errington RJ, Cook PR. *Embo J*. 1993; 12:1059. [PubMed: 8458323] (b) Wansink DG, Schul W, van der Kraan I, van Steensel B, van Driel R, de Jong L. *The Journal of cell biology*. 1993; 122:283. [PubMed: 8320255]
277. Iborra FJ, Pombo A, Jackson DA, Cook PR. *J Cell Sci*. 1996; 109(Pt 6):1427. [PubMed: 8799830]
278. Hozak P, Jackson DA, Cook PR. *J Cell Sci*. 1994; 107(Pt 8):2191. [PubMed: 7983177]
279. Melnik S, Deng B, Papantonis A, Baboo S, Carr IM, Cook PR. *Nat Methods*. 2011; 8:963. [PubMed: 21946667]
280. Mitchell JA, Fraser P. *Genes Dev*. 2008; 22:20. [PubMed: 18172162]
281. Ferrai C, Xie SQ, Luraghi P, Munari D, Ramirez F, Branco MR, Pombo A, Crippa MP. *PLoS Biol*. 2010; 8:e1000270. [PubMed: 20052287]
282. Larkin JD, Cook PR, Papantonis A. *Mol Cell Biol*. 2012; 32:2738. [PubMed: 22586273]

Biographies

David H. Price obtained his Ph.D in 1980 under the tutelage of Bill Marzluff at Florida State University, and then performed postdoctoral training under Eric Davidson and Carl Parker at Caltech. After working with Arno Greenleaf at Duke University as a Research Associate he took an Assistant Professor position in the Biochemistry Department at The University of Iowa where he has been a Full Professor since 1998. His research interests are in the control of human, *Drosophila*, and HIV gene expression. On most warm Sundays he can be found competing on an SCCA autocross course.

Jiannan Guo received his Bachelor's degree at the Northeast Forestry University in China and his Master's degree at the University of Bath in the UK. He then worked with Saverio Brogna at the University of Birmingham in the UK for his Ph.D. Currently, he is a postdoc in the Price lab at The University of Iowa in the US. Generally interested in the regulation of gene expression, his expertise includes developing in vitro transcription assays to characterize factors involved in Pol II transcription elongation.

Figure 1. Creating a nucleosome free region for PIC formation

Top, nucleosome occupancy inhibits promoter sequence recognition. **Middle**, occupancy of transcription factors (purple) and chromatin remodelers (pink) correlates with the nucleosome free region and promoter DNA is exposed. **Bottom**, GTFs, Pol II and Mediator form a PIC on the exposed promoter.

Figure 2. Comparison of sequencing based global methods

The diagrams illustrate the basic steps used in determining: **A)** Pol II or factor occupancy by ChIP-Seq, **B)** Identities and levels of specific RNAs in a population by RNA-Seq, and the locations of engaged Pol II utilizing **C)** nuclear run on (NRO) by GRO-Seq, or **D)** isolation of native elongating transcripts by NET-Seq.

Figure 3. Analysis of Pol II and selected transcription factors by ChIP-Seq

The ChIP-Seq datasets were obtained from published studies carried out by the Young lab on mouse embryonic stem cells.^{37b,69} Data were analyzed by Tiandao Li in the Price Lab using published methods.^{39a} **Top)** UCSC Genome Browser tracks showing ChIP-Seq results for Pol II, Pol II from cells treated with flavopiridol for 1 hr, Pol II with Ser5 or Ser2 CTD phosphorylation or the indicated factors over three chosen genes illustrate the range of Pol II occupancies found across the genome. For each protein, the heights of all three tracks were held constant and can, therefore, be compared across the different genes. **Bottom)** Metagene analysis showing the average signal for Pol II and the indicated factors from -600 to +600 from the TSS on 20,000 genes in mouse embryonic stem cells. Backgrounds were subtracted from the datasets before they were normalized.^{39a}

Figure 4. Promoter proximal paused Pol II

The diagram depicts Pol II that is paused under the control of factors including DSIF, NELF, Gdnaf1, the proposed Gdnaf1 Negative Accessory Factor (GNAF),^{39a} and TFIIS. TFIIF and TTF2 are blocked by Gdnaf1 at this stage, and Pol II is phosphorylated on Ser5 (illustrated as S5-P). The kinase module of Mediator, blocks entry of the next Pol II and facilitates CTD phosphorylation. The general activator and the acetylated chromatin binding protein BRD4 are also depicted although they are not active at this point.

Figure 5. The productive elongation complex

P-TEFb is recruited by BRD4 and triggers the transition into productive elongation. DSIF is phosphorylated turning it into a productive elongation factor and NELF is removed from the elongation complex. The CTD becomes phosphorylated on Ser2, which promotes mRNA processing. Components of SEC and Paf1C interact with both Pol II and Mediator. TFIIS facilitates release of the paused polymerase into productive elongation.

Figure 6. Reversible interaction of P-TEFb and HEXIM with the 7SK snRNP

The diagram depicts 7SK RNA (red line) with the indicated bound proteins before (**Top**) and after (**Bottom**) release of P-TEFb. The change in the shape of 7SK RNA represents the conformational change that occurs after release of P-TEFb which causes HEXIM to be ejected. BRD4 is known to cause the release of P-TEFb, but the re-sequestration of P-TEFb requires, as of yet, unknown activities.

Figure 7. Coupling of RNA processing with transcription and the influence of CTD phosphorylation

The diagram depicts Pol II at different stages of elongation across an average gene. The 7 amino acid repeat in the CTD is illustrated such that white circles are unphosphorylated and red circles are phosphorylated. Changes in the phosphorylation of the CTD and their influence on the indicated processing machineries are indicated.

Figure 8. Pol II termination pathways in yeast and human

The heptad repeat in the CTD is illustrated such that white circles are unphosphorylated and red circles are phosphorylated residues. Yeast Pol II can undergo termination mediated by the Nrd1/Nab3/Sen1 complex, which recognizes sequence elements in the transcript. In human, promoter proximal paused Pol II can be terminated by TTF2 unless Gdown1 is present. Once either yeast or human elongation complexes pass a Poly(A) signal and the nascent transcript is cleaved, termination is facilitated by the degradation of the nascent transcript by Rat1 in yeast or its homolog Xrn2 in human. A conformational change of the elongation complex and Pcf11 also promote termination.